



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 144864

TO: Jeffrey Parkin
Location: rem/3d39/3c18
Art Unit: 1648
Monday, February 28, 2005

Case Serial Number: 09/319156

From: Deirdre Arnold
Location: Biotech-Chem Library
REM 1A64
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

Please contact me if you encounter any problems with the disk or files.

Note that results are not generally sent via e-mail because the files are often very large.

Please feel free to contact me if you have any questions or would like to amend the search.

Thank you for using STIC services.

Regards,
Deirdre Arnold

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

144864

me

From: Parkin, Jeffrey
Sent: Wednesday, February 09, 2005 9:52 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 09/319,156

Please search SEQ ID NOS.: 6, 9, and 12 from the aforementioned application v. all relevant databases, including interference.

Place results on both PAPER and electronic format (i.e., e-mail).

Provide the first 40 results for each search.

Thanks!

JSP
AU 1648
REM 3D39
2-0908

RECEIVED
FEB 10 2005
STIC-Biotech/ChemLib
(STIC)

http://expoweb1:8001/cgi-bin/expo/GenInfo/snquery.pl?APPL_ID=09319156

STAFF USE ONLY

Searcher: Arnold
Searcher Phone: 2-2533
Date Searcher Picked up: 2/14/05
Date Completed: 2/28/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # 3
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Compu
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:32:54 ; Search time 2725.62 Seconds
(without alignments)
11288.850 Million cell updates/sec

Title: US-09-319-156B-6.
Perfect score: 635
Sequence: 1 cccgtgatttttaacctctc.....tgaaaaa.....635

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	635	100.0	635	6	BD136195 Retrovira
2	635	100.0	635	6	AR344385 Sequence
3	635	100.0	635	6	AX001024 Sequence
4	635	100.0	2004	14	AF127229 Multiple
5	631.8	99.5	167694	9	AC113137 Homo sapi
6	631.8	99.5	176095	2	AC019346 Homo sapi
7	631.8	99.5	210336	2	AC022171 Homo sapi
8	580.2	91.4	148724	2	AC010778 Homo sapi
9	579.2	91.2	8339	9	AL162912 Human DNA
10	571	89.9	136901	9	AC073626 Homo sapi
11	566.4	89.2	176188	9	AC068898 Homo sapi
12	564.6	88.9	89728	9	AL583805 Human DNA
13	560.2	88.2	112405	9	AC072023 Homo sapi
14	559.8	88.2	179114	9	BS000045 Pan trogl
15	556.6	87.7	174019	9	AP001538 Homo sapi
16	556.6	87.7	340000	2	AP001674 Homo sapi
17	551.8	86.9	149755	2	AP001545 Homo sapi
18	551.8	86.9	152980	2	AC090313 Homo sapi
19	551.8	86.9	163803	9	AC093531 Homo sapi

C 20	551.8	86.9	176425	9	AC107075	AC107075 Homo sapi
C 21	551.4	86.8	128468	9	AL139090	AL139090 Human DNA
C 22	551.4	86.8	140756	9	AL139038	AL139038 Human DNA
C 23	551.4	86.8	167366	2	AC021774	AC021774 Homo sapi
C 24	551.4	86.8	183499	2	AL607153	AL607153 Homo sapi
C 25	550.2	86.6	164211	2	AP002790	AP002790 Homo sapi
C 26	550.2	86.6	167843	9	AC064801	AC064801 Homo sapi
C 27	549	86.5	180523	9	CNS01DSH	AL121784 Human chr
C 28	549	86.5	190565	9	AC007374	AC007374 Homo sapi
C 29	549	86.5	203777	9	CNS01DVH	AL135818 Human chr
C 30	547.8	86.3	112544	9	AC109992	AC109992 Homo sapi
C 31	547.6	86.2	192178	2	AC009443	AC009443 Homo sapi
C 32	547.6	86.2	192899	9	CNS05TF5	AL352955 Human chr
C 33	547.4	86.2	172281	9	AC068492	AC068492 Homo sapi
C 34	547.2	86.2	169418	9	AC104163	AC104163 Homo sapi
C 35	546.6	86.1	161054	9	AC104444	AC104444 Homo sapi
C 36	546.4	86.0	2074	6	AX478550	AX478550 Sequence
C 37	546	86.0	169234	9	CNS06C7Q	AL390799 Human chr
C 38	545	85.8	1329	6	BD136199	BD136199 Retrovira
C 39	545	85.8	1329	6	AR344389	AR344389 Sequence
C 40	545	85.8	1329	6	AX001030	AX001030 Sequence
C 41	545	85.8	99408	9	AL359385	AL359385 Human DNA
C 42	545	85.8	170586	2	AL356126	AL356126 Homo sapi
C 43	543.8	85.6	181166	9	AC087897	AC087897 Homo sapi
C 44	543.4	85.6	88328	9	AL357874	AL357874 Human DNA
C 45	539	84.9	2082	2	AC134379	AC134379 Homo sapi

ALIGNMENTS

RESULT 1	BD136195	635 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD136195				
DEFINITION	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.				
ACCESSION	BD136195				
VERSION	BD136195.1	GI:232331140			
KEYWORDS	JP 2002509437-A/5.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 635)				
AUTHORS	Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F., Perron,H. and Mandrand,B.				
TITLE	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses				
JOURNAL	Patent: JP 2002509437-A 5 26-MAR-2002;				
COMMENT	BIO MERIEUX				
	OS Unidentified				
	PN JP 2002509437-A/5				
	PD 26-MAR-2002				
	PF 07-JUL-1998 JP 1999508255				
	PR 07-JUL-1997 FR 97/08816				
	PI GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC PI BEDIN				
	PI MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON, PI BERNARD MANDRAND				
	PC C12N15/48,C12Q1/70,C07K14/15,A61K31/70				
	CC Strandedness: Single;				
	CC Topology: Linear;				
	CC Retroviral nucleic material and nucleotide fragments, in CC particular,				
	CC associated with multiple sclerosis and/or				
	rheumatoid arthritis,				
	CC for				
	CC diagnostic, prophylactic and therapeutic uses FH Key				
	Location/Qualifiers				
FT	source	1..635			
FT	/organism='Unidentified'.				
	Location/Qualifiers				

source		1. .635		/organism="unidentified"	
				/mol_type="genomic DNA"	
				/db_xref="taxon:32644"	
ORIGIN					
Query Match		100.0%; Score 635; DB 6; Length 635;			
Best Local Similarity		100.0%; Pred. No. 1.1e-189;			
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60		
Db	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60		
Qy	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120		
Db	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120		
Qy	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180		
Db	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180		
Qy	181	GAATCTCAACTGCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240		
Db	181	GAATCTCAACTGCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240		
Qy	241	CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTGAGAGGGGGGACTGAGAGAC	300		
Db	241	CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTGAGAGGGGGGACTGAGAGAC	300		
Qy	301	AGGACTAGCTGGAATTTCTTAGCCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360		
Db	301	AGGACTAGCTGGAATTTCTTAGCCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360		
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420		
Db	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420		
Qy	421	ACTAAATGCTAATAGGCAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480		
Db	421	ACTAAATGCTAATAGGCAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480		
Qy	481	AGAGCACAGCGGAGGACCAAGATCGGATATAAACCCAGGCATTCGAGCCGCAACGG	540		
Db	481	AGAGCACAGCGGAGGACCAAGATCGGATATAAACCCAGGCATTCGAGCCGCAACGG	540		
Qy	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600		
Db	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600		
Qy	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635		
Db	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635		
RESULT 2					
AR344385		635 bp		linear	
LOCUS					
DEFINITION		Sequence 102 from patent US 6582703.		PAT 17-AUG-2003	
ACCESSION		AR344385			
VERSION		AR344385.1		GI:33740326	
KEYWORDS					
SOURCE		Unknown.			
ORGANISM		Unclassified.			
REFERENCE		1 (bases 1 to 635)			
AUTHORS		Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G., Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.			
TITLE		Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting			
JOURNAL		Patent: US 6582703-A 102 24-JUN-2003;			
FEATURES		Location/Qualifiers			
source		1. .635		/organism="unidentified"	
ORIGIN					
Query Match		100.0%; Score 635; DB 6; Length 635;			
Best Local Similarity		100.0%; Pred. No. 1.1e-189;			
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60		
Db	1	CCCTGTATCTTTAAACCTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA	60		
Qy	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120		
Db	61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120		
Qy	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180		
Db	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGAG	180		
Qy	181	GAATCTCAACTGCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240		
Db	181	GAATCTCAACTGCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT	240		
Qy	241	CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTGAGAGGGGGGACTGAGAGAC	300		
Db	241	CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTGAGAGGGGGGACTGAGAGAC	300		
Qy	301	AGGACTAGCTGGAATTTCTTAGCCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360		
Db	301	AGGACTAGCTGGAATTTCTTAGCCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT	360		
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420		
Db	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420		
Qy	421	ACTAAATGCTAATAGGCAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480		
Db	421	ACTAAATGCTAATAGGCAAAATAGGAGTTAAGAAATAGCCATCATCTATTGCCTG	480		
Qy	481	AGAGCACAGCGGAGGACCAAGATCGGATATAAACCCAGGCATTCGAGCCGCAACGG	540		
Db	481	AGAGCACAGCGGAGGACCAAGATCGGATATAAACCCAGGCATTCGAGCCGCAACGG	540		
Qy	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600		
Db	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCACTCT	600		
Qy	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635		
Db	601	ATTAATCTTGCNACTGAAAAAAGAAAAA	635		
RESULT 3					
AX001024		635 bp		linear	
LOCUS					
DEFINITION		Sequence 6 from Patent WO9902666.		PAT 10-MAR-2000	
ACCESSION		AX001024			
VERSION		AX001024.1		GI:7241262	
KEYWORDS					
SOURCE		unidentified			
ORGANISM		unclassified.			
REFERENCE		1 (bases 1 to 635)			
AUTHORS		Ott,C. and Bedin,F.			
TITLE		RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES			
JOURNAL		Patent: WO 9902666-A 6 21-JAN-1999;			
FEATURES		Location/Qualifiers			
source		1. .635		/organism="unidentified"	

REFERENCE		2 (bases 1 to 2004)	
AUTHORS		Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F., Ounian-Paraz, A., Sodoyer, M., Ott, C., Rajoharison, A., Garcia, E., Mallet, F., Mandrand, B. and Perron, H.	
TITLE		Direct Submission	
JOURNAL		Submitted (10-FEB-1999) UMR103 CNRS, bioMerieux, 46, allée D'Italie, Lyon 69007, France	
FEATURES		Location/Qualifiers	
source		1..2004	
		/organism="Multiple sclerosis associated retrovirus element"	
		/mol_type="mRNA"	
		/db_xref="taxon:89382"	
		/clone="CL6"	
misc_feature		1..2004	
		/note="similar to pol-env and 3'LTR region"	
ORIGIN			
Query Match		100.0%; Score 635; DB 14; Length 2004;	
Best Local Similarity		100.0%; Pred. No. 1.3e-189;	
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CCCTGTATCTTTAACTCCTCTGTTAAGTTTGTCTCTCCAGATCAAACTGTAAACTA	60
DB	1370	CCCTGTATCTTTAACTCCTCTGTTAAGTTTGTCTCTCCAGATCAAACTGTAAACTA	1429
QY	61	CAAAATTTGTTCTTCAATGGAGCAGATGGAGTCCATGACTAAGATCCACCGTGACCC	120
DB	1430	CAAAATTTGTTCTTCAATGGAGCAGATGGAGTCCATGACTAAGATCCACCGTGACCC	1489
QY	121	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACCCCTCCCGAG	180
DB	1490	CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGGACCCCTCCCGAG	1549
QY	181	GAATCTCACTGACGACACCCCTACTATGCTCCCAATTCAGCGGAGCAGTTAGCGGT	240
DB	1550	GAATCTCACTGACGACACCCCTACTATGCTCCCAATTCAGCGGAGCAGTTAGCGGT	1609
QY	241	CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTCTGTAGAGGGGGGAGCTGAGAG	300
DB	1610	CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTCTGTAGAGGGGGGAGCTGAGAG	1669
QY	301	AGGACTAGCTGGATTTCCTAGCCCAACGAAGATCCCTAAGCTAGCTGGGAGGTGACT	360
DB	1670	AGGACTAGCTGGATTTCCTAGCCCAACGAAGATCCCTAAGCTAGCTGGGAGGTGACT	1729
QY	361	GCATCCACCTCTAAACATGGGCTTGCATCTAGCTCACACCGGACCAATCAGAGCTC	420
DB	1730	GCATCCACCTCTAAACATGGGCTTGCATCTAGCTCACACCGGACCAATCAGAGCTC	1789
QY	421	ACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG	480
DB	1790	ACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG	1849
QY	481	AGAGCAGCGGGAGGACCAAGGATCGGATATAAACCAGGCAATTCGAGCCGGCAACGG	540
DB	1850	AGAGCAGCGGGAGGACCAAGGATCGGATATAAACCAGGCAATTCGAGCCGGCAACGG	1909
QY	541	CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACCT	600
DB	1910	CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACCT	1969
QY	601	ATTAAATCTTGCAACTGAAAAAAGAAAAAAGAAAAA 635	
DB	1970	ATTAAATCTTGCAACTGAAAAAAGAAAAAAGAAAAA 2004	
RESULT 5			
AC113137/c		167694 bp DNA linear PRI 25-APR-2002	
LOCUS		Homo sapiens chromosome 18, clone RP11-622J9, complete sequence..	
DEFINITION		AC113137	
ACCESSION		AC113137	
VERSION		AC113137.2	
PUBMED		GI:19683504	

HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-622J9
Unpublished
2 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Farre,S., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 167694)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Farreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,<

Db	123369	CATCAGCCAACTCC	CCCAACAGC	CAC	TGGGTTTTCTCTGTTGAGAGGGGGAGCTGAGAGAC	123311
QY	301	AGACTAGCTGGATTT	CCTAGGCCA	AGAGAA	TCCCTTAAGCCCTAGCTGGGAAGGTGACT	360
Db	123309	AGGACTAGCTGGATTT	CCTAGGCCA	AGAGAA	TCCCTTAAGCCCTAGCTGGGAAGGTGACT	123250
QY	361	GCATCCACCTCTAAA	CATATGGGG	TTGCA	ACTTAGCTCACACCCGACCAATCAGAGAGCTC	420
Db	123249	GCATCCACCTCTAAA	CATATGGGG	TTGCA	ACTTAGCTCACACCCGACCAATCAGAGAGCTC	123190
QY	421	ACTAAATGCTAAT	TAGGC	AAAAA	TATAGGAGTAAAGAAATAGCCAATCATCTATTGCTG	480
Db	123189	ACTAAATGCTAAT	TAGGC	AAAAA	TATAGGAGTAAAGAAATAGCCAATCATCTATTGCTG	123130
QY	481	AGAGCAGCGGAGGGA	CAAGATCGG	ATATAA	CCAGGCAATCGAGCGCGCAACGG	540
Db	123129	AGAGCAGCGGAGGGA	CAAGATCGG	ATATAA	CCAGGCAATCGAGCGCGCAACGG	123070
QY	541	CAACCCCTTTGGT	CCCCCTCC	CTTTCTAT	TGGGGCTCTGTTTTCACCTATTTCACCTCT	600
Db	123069	CAACCCCTTTGGT	CCCCCTCC	CTTTCTAT	TGGGGCTCTGTTTTCACCTATTTCACCTCT	123010
QY	601	ATTAATCTTGC	AACTG	AAAAA	AAAAA	635
Db	123009	ATTAATCTTGC	AACTG	AAAAA	AAAAA	122975
RESULT 6						
AC019346/c						
LOCUS			176095 bp	DNA	linear	HTG 25-OCT-2001
DEFINITION						Homo sapiens chromosome 18 clone RP11-497M7 map 18, WORKING DRAFT SEQUENCE.
ACCESSION						AC019346
VERSION						GI:16418201
KEYWORDS						HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE						Homo sapiens (human)
ORGANISM						Homo sapiens
REFERENCE						Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS						1 (bases 1 to 176095)
TITLE						Bliren,B., Linton,L., Nusbaum,C. and Landet,E.
JOURNAL						Homo sapiens chromosome 18, clone RP11-497M7
REFERENCE						Unpublished
AUTHORS						2 (bases 1 to 176095)
						Bliren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boquelavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardaya,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
TITLE						Direct Submission
JOURNAL						Submitted (02-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT						On Oct 25, 2001 this sequence version replaced gi:7230183. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
						----- Genome Center
						Center: Whitehead Institute/ MIT Center for Genome Research
						Center code: WITB
						Web site: http://www-seq.wi.mit.edu
						Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L4296
Center clone name: 497_M7

----- Summary Statistics

Sequencing vector: M13; M77815; 32% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 176087 bases at least Q40
Consensus quality: 176095 bases at least Q30
Insert size: 178000; agarose-fp
Quality coverage: 11.5 in Q20 bases; agarose-fp
Quality coverage: 11.7 in Q20.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 176095: contig of 176095 bp in length.

FEATURES

source

1. 176095
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-497M7"
/clone_lib="RPC1-11 Human Male BAC"

ORIGIN

Query Match 99.5%; Score 631.8; DB 2; Length 176095;
Best Local Similarity 99.7%; Pred No. 2.9e-188;
Matches 633; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACCTCTGTTAAAGTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60
DB 150524 CCTGTATCTTTAACCTCTGTTAAAGTTGTCTCTCCAGAAATCAAAACTGTAAACTA 150465

QY 61 CAAATGTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 150464 CAAATGTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 150405

QY 121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGATTTGAAGGCAACCCCTCCCGAG 180
DB 150404 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGATTTGAAGGCAACCCCTCCCGAG 150345

QY 181 GAAATCTCAATGTCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 150344 GAAATCTCAATGTCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 150285

QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTCTCTTGGAGGGGGGAGTGAAGAC 300
DB 150284 CATCAGCCAACTCCCAACAGCACTTGGGTTTCTCTTGGAGGGGGGAGTGAAGAC 150225

QY 301 AGGACTAGCTGGATTTCCTAGGCCAAACGAAGAAATCCCTAAGCTAGCTGGGAAGGTGACT 360
DB 150224 AGGACTAGCTGGATTTCCTAGGCCAAACGAAGAAATCCCTAAGCTAGCTGGGAAGGTGACT 150165

QY 361 GCATCCACTCTAAACATGGGGCTTGGAACTTAGTCTACACCCGACCAATCAGAGAGCTC 420
DB 150164 GCATCCACTCTAAACATGGGGCTTGGAACTTAGTCTACACCCGACCAATCAGAGAGCTC 150105

QY 421 ACTAAATGCTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCTGT 480
DB 150104 ACTAAATGCTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCTGT 150045

QY 481 AGACACAGCGGAGGACAGGATCGGATATAAACCCAGCATTCGAGCGGCAACGG 540
DB 150044 AGACACAGCGGAGGACAGGATCGGATATAAACCCAGCATTCGAGCGGCAACGG 149985

QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCTCTCT 600
DB 149984 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTCTCTCT 149925

QY 601 ATTAAATCTTGCACACTGAAAAAAGAAAAA 635
DB 149924 ATTAAATCTTGCACACTGAAAAAAGAAAAA 149890

RESULT 7

AC022171

LOCUS

DEFINITION

AC022171

AC022171

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Mar 10, 2001 this sequence version replaced gi:13122770.

----- Genome Center

Center: Stanford DNA Sequencing and Technology Development

Center

Center code: SDSTDC

Web site: <http://sequence-www.stanford.edu/group/human/>

Contact: hum-info@sequence.stanford.edu

----- Project Information

Center project name: 758

Center clone name: RP11-407C18

----- Summary Statistics

Sequencing Vector: M13mpl8; X02513; 97% of reads

Sequencing Vector: plasmid; X02513; 97% of reads

Chemistry: Dye-terminator Big Dye; 88% of reads

Chemistry: Dye-terminator Big Dye; 88% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 208312 bases at least Q40

Consensus quality: 208877 bases at least Q30

Consensus quality: 209157 bases at least Q20

Insert size: 199762; agarose-fp

Insert size: 210236; sum-of-contigs

Quality coverage: 8.7x in Q20 bases; agarose-fp

Quality coverage: 8.3x in Q20 bases; sum-of-contigs.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 66468: contig of 66468 bp in length

* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

Fri Feb 25 16:26:29 2005

```
* 98588 98687: gap of 100 bp
* 98688 148724: contig of 50037 bp in length.
FEATURES
    Location/Qualifiers
        1..148724
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="RP11-2N15"
            /clone_lib="RPC1-11 Human Male BAC"
        1..2462
            /note="assembly_fragment"
        2563..7948
            /note="assembly_fragment"
        8049..15800
            /note="assembly_fragment"
            clone_end:77
            vector_side:right
        15901..25059
            /note="assembly_fragment"
        25160..41906
            /note="assembly_fragment"
        42007..63341
            /note="assembly_fragment"
        63442..98597
            /note="assembly_fragment"
        98688..148724
            /note="assembly_fragment"
    ORIGIN
        Query Match          91.4%;   Score 580.2;   DB 2;   Length 148724;
        Best Local Similarity 94.8%;   Pred. No. 6.6e-172;
        Matches 600;   Conservative 0;   Mismatches 33;   Indels 0;   Gaps 0;
QY 1 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
DB 40271 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTTCAGAAATGAAACTGTAAACTA 40212
QY 61 CAAATTGTTCTTCAATGAGCACGATGAGTCCATGATCAAGTCCACCGTGAGCC 120
DB 40211 CAAATTGTTCTTCAATGAGCCCCAGATGCGATGCATGATCAAGTCCACCGTGAGTCC 40152
QY 121 CTGACCGCGCTGTAGCCATGCTCCGATGTTAATGACATTTGAAGCACCCCTCCCGAG 180
DB 40151 CTGACCGCGCTGTAGCCATGCTCCGATGTTAATGATCGAAGGACCCCTCCCGAG 40092
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTTAGACGGT 240
DB 40091 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTTAGACGAT 40032
QY 241 CATCAGCCCACTCCCAACGACATTTGGTTTCTGTTGAGAGGGGACTTGAGAGAC 300
DB 40031 CATCAGCCCACTCCCAACGACATTTGGTTTCTGTTGAGAGGGGACTTGAGAGAC 39972
QY 301 AGGACTAGCTGGATTTTCTAGGCCAACGAGAAATCCCTAAGCCTAGCTGGGAGGTGACT 360
DB 39971 AGGACTAGCTGGATTTTCTAGGCCAACGAGAAATCCCTAAGCCTAGCTGGGAGGTGACT 39912
QY 361 GCATCCACCTCTTAAACATGCGGCTTGCACTTAGCTTCACACCGGACCAATCAGAGCTC 420
DB 39911 ACACCCACCTTTAAACATGCGGCTTGCACTTAGCTTCACACCCCAACCAATCAGAGCTC 39852
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 480
DB 39851 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 39792
QY 481 AGAGCAGCGGGGAGGACAGGATCGGGATATAAACCAGGCAATTCGAGCCGGCAACGG 540
DB 39791 AGAGCAGCTGGAGGGACAGGATCGGGATATAAACCAGGCAATTCGAGCCAGCAACAG 39732
QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCCTCT 600
DB 39731 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCCTCT 39672
```

```
QY 601 ATTAAATCTTGCAACTGAAAAAAGAAAAA 633
DB 39671 ATTAAATCTTGCAACTGAAAAAAGAAAAAATGCAA 39639

RESULT 9
AL162912
LOCUS
DEFINITION
    Human DNA sequence from clone RP6-198C4 on chromosome Xq13.3-21.2,
    complete sequence.
ACCESSION
    AL162912 AL121824
VERSION
    AL162912.1 GI:7406722
KEYWORDS
    HTG.
SOURCE
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 8339)
    Pearce,A.
    Direct Submission
    Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
    Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
    humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
    During sequence assembly data is compared from overlapping clones.
    Where differences are found these are annotated as variations
    together with a note of the overlapping clone name. Note that the
    variation annotation may not be found in the sequence submission
    corresponding to the overlapping clone, as we submit sequences with
    only a small overlap as described above.
    The following abbreviations are used to associate primary accession
    numbers given in the feature table with their source databases:
    Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
    on the WORMPEP database can be found at
    http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
    was generated from part of bacterial clone contigs of human
    chromosome X, constructed by the Sanger Centre Chromosome X Mapping
    Group. Further information can be found at
    http://www.sanger.ac.uk/HGP/ChrX
    This sequence was finished as follows unless otherwise noted: all
    regions were either double-stranded or sequenced with an alternate
    chemistry or covered by high quality data (i.e., phred quality >=
    30); an attempt was made to resolve all sequencing problems, such
    as compressions and repeats; all regions were covered by at least
    one plasmid subclone or more than one M13 subclone; and the
    assembly was confirmed by restriction digest. RP6-198C4 is from the
    library RPC1-6 constructed by the group of Pieter de Jong. For
    further details see
    http://www.chori.org/bacpac/home.htm
    VECTOR: pPAC4
    IMPORTANT: This sequence is not the entire insert of clone
    RP6-198C4 it may be shorter because we sequence overlapping
    sections only once, except for a short overlap.
FEATURES
    Location/Qualifiers
        1..8339
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="X"
            /map="q13.3-21.2"
            /clone="RP6-198C4"
            /clone_lib="RPC1-6"
        1..5641
            /note="HERV17 repeat: matches 368..5930 of consensus"
        5635..5673
            /note="HERV17 repeat: matches 7419..7457 of consensus"
        5664..5955
            /note="HERV17 repeat: matches 8244..8523 of consensus"
        5956..6136
            /note="L1R17 repeat: matches 1..239 of consensus"
        6195..6280
            /note="L1R17 repeat: matches 239..324 of consensus"
    ORIGIN
```

```

Query Match          91.2%; Score 579.2; DB 9; Length 8339;
Best Local Similarity 94.8%; Pred. No. 8 4e-172;
Matches 599; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTTTCAGAAATCAAACTGTAATAACTA 60
Db CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTTTCAGAAATCAAACTGTAATAACTA 5723

QY 61 CAAATGTTCTTCAATGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db CAAATGTTCTTCAATGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGATCC 5783

QY 121 CTGACCGGCTCTAGCCATCTCCGATGTTAAATGACATTAAGGACACCCCTCCCGAG 180
Db CTGACCGGCTCTAGCCATCTCCGATGTTAAATGACATCGAAGGACACCCCTCCCGAG 5843

QY 181 GAAATCTCAATGCACAAACCCCTTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACGGT 240
Db GAAATCTCAATGCACAAACCCCTTACTATGCCCCCAATTCAGCAGGAAGCAGTTAGACGAT 5903

QY 241 CATCAGCAACCTCCCAACAGCAGCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db CATCAGCAACCTCCCAACAGCAGCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 5963

QY 301 AGGACTAGCTGGATTTCTAGGCAAGAGATCCTTAAGCTAGCTGGGAAGGTGACT 360
Db AGGACTAGCTGGATTTCTAGGCAAGAGATCCTTAAGCTAGCTGGGAAGGTGACT 6023

QY 361 GCATCCACCTCTAAACATGSGGCTTCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db GCATCCACCTCTAAACATGSGGCTTCACTTAGCTCACACCCGACCAATCAGAGAGCTC 6083

QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTCCCTG 480
Db ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTCCCTG 6143

QY 481 AGAGCAGCGGAGGAGCAGAGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 540
Db AGAGCAGCGGAGGAGCAGAGATCGGATATAAACCCAGGATTCGAGCCGGCAACAG 6203

QY 541 CAACCCCTCTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCCTCT 600
Db CAACCCCTCTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCCTCT 6263

QY 601 ATTAATCTTGCACGTAAGAAAAA 632
Db ATTAATCTTGCACGTAAGAAAAAATGCAA 6295

```

```

RESULT 10
AC073626/c 136901 bp DNA linear PRI 30-JAN-2004
LOCUS Homo sapiens BAC clone RP11-95P9 from 7, complete sequence.
DEFINITION AC073626
ACCESSION AC073626.7 GI:12863221
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS Hillier, L.W., Fulton, R.S., Fulton, L.A., Graves, T.A., Pepin, K.H.,
Wagner-McPherson, C., Layman, D., Maas, J., Jaeger, S., Walker, R.,
Wylie, K., Sekhon, M., Becker, M.C., O'Laughlin, M.D., Schaller, M.E.,
Fowell, G.A., Delehaunty, K.D., Miner, T.L., Nash, W.E., Cordes, M.,
Du, H., Sun, H., Edwards, J., Bradshaw-Cordum, H., Ali, J., Andrews, S.,
Isak, A., Vanbrunt, A., Nguyen, C., Du, F., Lamar, B., Courtney, L.,
Kalicki, J., Ozersky, P., Bledick, L., Scott, K., Holmes, A.,
Harkins, R., Harris, A., Strong, C.M., Hou, S., Tomlinson, C.,
Dauphin-Kohlberg, S., Kozlowicz-Reilly, A., Leonard, S., Rohlfing, T.,
Rock, S.M., Tin-Wollam, A.M., Abbott, A., Minx, P., Maupin, R.,
Stromwatt, C., Latreille, P., Miller, N., Johnson, D., Murray, J.,
Weesener, J.P., Wendt, M.C., Yang, S.P., Schultz, B.R., Wallis, J.W.,

```

Spith, J., Bieri, T.A., Nelson, J.O., Berkowicz, N., Wohldmann, P.E., Cook, L.L., Hickenbotham, M.T., Eldred, J., Williams, D., Bedell, J.A., Mardis, E.R., Clifton, S.W., Chissoe, S.L., Marra, M.A., Raymond, C., Haugen, E., Gillett, W., Zhou, Y., James, R., Phelps, K., Iadamoto, S., Bub, K., Simms, E., Levy, R., Clendenning, J., Kaul, R., Kent, W.J., Furey, T.S., Baerbach, R.A., Brent, M.R., Keibler, E., Flicek, P., Bork, P., Suyama, M., Bailey, J.A., Portnoy, M.E., Torrens, D., Chinwalla, A.T., Gish, W.R., Eddy, S.R., McPherson, J.D., Olson, M.V., Eichler, E.E., Green, E.D., Waterston, R.H. and Wilson, R.K.

The DNA sequence of human chromosome 7
Nature 424 (6945), 157-164 (2003)
22737999

MEDLINE
PUBMED
12853948
2 (bases 1 to 136901)
The sequence of Homo sapiens BAC clone RP11-95P9

AUTHORS
Bauer, H., Haakenson, B. and Nguyen, C.
TITLE
Unpublished (2001)

JOURNAL
3 (bases 1 to 136901)
Waterston, R.H.
Direct Submission

AUTHORS
Submitted (27-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

JOURNAL
4 (bases 1 to 136901)
Waterston, R.H.
Direct Submission

AUTHORS
Submitted (16-FEB-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

JOURNAL
5 (bases 1 to 136901)
Waterston, R.H.
Direct Submission

AUTHORS
Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

JOURNAL
6 (bases 1 to 136901)
Waterston, R.H.
Direct Submission

AUTHORS
Submitted (29-APR-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

JOURNAL
7 (bases 1 to 136901)
Wilson, R.
Direct Submission

AUTHORS
Submitted (30-JAN-2004) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

JOURNAL
On Feb 16, 2001 this sequence version replaced gi:11597125.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_NH0095P09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information

SOURCE INFORMATION:
The RPlci-11 human BAC library was made from the blood of one male donor, as described by Onoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tatenom, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Piter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is CTD-2023N18, 200 bp overlap the
clone sequenced to the right is CTA-250D13, 200 bp overlap. Actual
start of this clone is at base position 137066 of CTD-2023N18
and actual end is at base position 26937 of CTA-250D13.

repeat_region	545..1149
rpt_family="L1"	/rpt_family="L1"
repeat_region	1195..1380
rpt_family="L1"	/rpt_family="L1"
repeat_region	1533..1790
rpt_family="L1"	/rpt_family="L1"
repeat_region	2972..3112
rpt_family="MERI_type"	/rpt_family="MERI_type"
repeat_region	3118..3324
rpt_family="MERI_type"	/rpt_family="MERI_type"
repeat_region	6682..6885
rpt_family="MIR"	/rpt_family="MIR"
repeat_region	7038..7388
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	7433..7770
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	7786..7823
rpt_family="AT_rich"	/rpt_family="AT_rich"
repeat_region	7829..7985
rpt_family="(TA)n"	/rpt_family="(TA)n"
repeat_region	8004..8064
rpt_family="(TATG)n"	/rpt_family="(TATG)n"
repeat_region	8067..8206
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	9047..9246
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	9247..9288
rpt_family="(TCCTCG)n"	/rpt_family="(TCCTCG)n"
repeat_region	9289..9399
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	9412..9494
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	9495..9854
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	9855..9928
rpt_family="MaLR"	/rpt_family="MaLR"
repeat_region	10530..10814
rpt_family="Alu"	/rpt_family="Alu"
repeat_region	10815..10909
rpt_family="(CATA)n"	/rpt_family="(CATA)n"
repeat_region	14398..14434
rpt_family="(TAA)n"	/rpt_family="(TAA)n"
repeat_region	14453..14659
rpt_family="MIR"	/rpt_family="MIR"
repeat_region	15191..15248

```
repeat_region    /rpt_family="Achobo"  
15255. .15456  
/rpt_family="MaLR"  
15457. .15494  
/rpt_family="AT_rich"  
15515. .15552  
/rpt_family="AT_rich"  
15997.16043  
/rpt_family="(CA)n"  
16497. .18517  
/rpt_family="AT_rich"  
16886. .17023  
/rpt_family="MERI_type"  
17189. .17380  
/rpt_family="MERI_type"  
17513. .17684  
/rpt_family="MIR"  
18387.18555  
/rpt_family="MIR"  
18604. .19120  
/rpt_family="MaLR"  
20362. .20667  
/rpt_family="Alu"  
20943. .20966  
/rpt_family="AT_rich"  
21716. 21879  
/rpt_family="MIR"  
22121. .22550  
/rpt_family="ERVL"  
22576. .23085  
/rpt_family="MaLR"  
23097. .24931  
repeat_region
```

Query Match	89.9%;	Score 571;	DB 9;	Length 136901;
Best Local Similarity	93.7%;	Pred. No. 5.4e-169;		
Matches 595;	Conservative	0;	Mismatches 40;	Indels 0; Gaps 0;
QY	1	CCCTGTATCTTTAACTCCTCTTGTAAAGTTTGTCTCTTCCAGAATCAAAACTGTAAAACTA	60	
DB	95677	CCCTGTATCTTTAACTCCTCTTGTAAAGTTTGTCTCTTCCAGAATTGAACACTGTAAAACTA	95618	
QY	61	CAAAATTGTTCTTTCAAAATGGAGCACACAGATGGAGTCCATGACTAGATCCACCGTGGACCC	120	
DB	95617	CAAAATGGTTTTTCAAAATGGAGCCCAGATGCAGTCCATGACTAAAGATCCACCGTGGACCC	95558	
QY	121	CTGGACCGGCGCTTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCACACCCCTCCCGAG	180	
DB	95557	CTGAATGGCGCTGCTAGCCCATTCCTCATGTGTTAATGATGTGGAAGGCACCCCTCCCGAG	95498	
QY	181	GAAACTCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGGAAGCAGTTAGAGCGGT	240	
DB	95497	AAAACTCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCAGGAAGCAGTTAGAGTGGT	95438	
QY	241	CATCAGCCAACTCCGCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC	300	
DB	95437	CATCGGCCAACTCCGCCAACAACTTGGGTTTTCTGTTGAGAGGGGGTACTGAGAGAC	95378	
QY	301	AGGACTAGCTGGATTTCTTAGCCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT	360	
DB	95377	AGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGNAAGTGACC	95318	
QY	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTTCACACCCGACCAATCAGAGAGCTC	420	
DB	95317	GCATCCACCTTTAAACATGGTGCTCGCACTTAGCTCCACCACCAATCAGAGAGCTC	95258	
QY	421	ACTAAAATGCTAATTAGGCCAAAAATAGGAGTTAAAGAAATGCCAATCATCTATTGGCTTG	480	
DB	95257	ACTAAAATGCTAATTATGCAAAAAAGGAGGTAAAGAAATAGCCCAATCATCTATTGGCTTG	95198	
QY	481	AGAGCACAGCGGAGGGACAAGGATCGGATATAACCCAGGCAATTCAGAGCCGCAACGG	540	
DB	95197	AGAGCACAGGTGGAGGGACAAGGATCGGACATAAACCCAGGCAATTCAGAGCCAGCAACAG	95138	

```

Qy 541 CRACCCCTTTGGTCCCTCCCTTTGTTATGGCGCTCTGTTTCACTCTATTCTACTCT 600
Db 95137 CRACCCCTTTGGTCCCTCCCTCAATTTGTTATGGAGCTCTGTTTCACTCTATTCTACTCT 95078
Qy 601 ATTAATCTTGGCACTGAAAAAATAAAAAA 635
Db 95077 ATTAATCTTGGCACTGCAAAAAATAAAAAA 95043

RESULT 11
AC068898/c 176188 bp DNA linear PRI 29-AUG-2002
LOCUS Homo sapiens chromosome 10 clone RP11-534L6, complete sequence.
AC068898
AC068898.8 GI:22539070
VERSION HTG.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 176188)
Smith, D.R.
Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 176188)
Smith, D.R.
Direct Submission
JOURNAL Submitted (11-MAY-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 3 (bases 1 to 176188)
Smith, D.R.
Direct Submission
JOURNAL Submitted (29-AUG-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Aug 29, 2002 this sequence version replaced gi:14625082.
FEATURES
source
1..176188
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-534L6"
/clone_lib="RPCI-11"

ORIGIN
Query Match 89.2%; Score 566.4; DB 9; Length 176188;
Best Local Similarity 93.5%; Pred. No. 1.6e-167;
Matches 591; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 1 CCTGTATCTTTAACTCCTTGTAGTTTGTCTTCCAGATCAAACTGTAAACTA 60
Db 57896 CCCTGTATCTTTAACTCCTTGTAGTTTGTCTTCCAGATCAAACTGTAAACTA 57837

Qy 61 CAAATTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAGATCCACCGTGGACCC 120
Db 57836 CAAATTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAGATCCACCGTGGACCC 57777

Qy 121 CTGACCGGCTGTAGCCCATGCTCGATGTTTAATGACATTTGAAGGACCCCTCCCGAG 180
Db 57776 CTAGACTGGCTGTAGCCCATGCTCGATGTTTAATGACATTTGAAGGACCCCTCCCGAG 57717

Qy 181 GAAATCTCAATGCAACCCCTACTATGCCCATTACGGGGGAGCAGTTAGAGCGGT 240
Db 57716 GAAATCTCAATGCAACCCCTACTATGCCCATTACGGGGGAGCAGTTAGAGCGGT 57657

Qy 241 CATCAGCCACCTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 57656 CATCAGCCACCTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGAGCTGAGAGAC 57597

Qy 301 AGGACTAGCTGGATTCTCTAGGCCAACAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACT 360
Db 57596 AGGACTAGCTGGATTCTCTAGGCCGATTAAAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACC 57537

```

```

Qy 361 GCATCCACCTCTAAACATGGGGCTTGCACACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 57536 GCATCCACCTTTAAACACGGGGCTTGCACACTTAGCTCAACCCGACCAATCAGAGAGCTC 57477
Qy 421 ACTAAATGCTAATTAGGCAAAATAGAGGTAAGAAATAGCCATCATCTATTGCTCG 480
Db 57476 ACTAAATGCTAATTAGGCAAAACAGGAGGTAAGAAATAGCCATCATCTATTGCTCG 57417
Qy 481 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCGCGCAACGG 540
Db 57416 AGAGCACAGTGGAGGAGGACAGGATTCGATATAAACCCAGGATTCGAGCGCGCAACGG 57357
Qy 541 CAACCCCTTTGGTCCCTCCCTTTGTTATGGCGCTCTGTTTCACTCTATTCTACTCT 600
Db 57356 CAACCCCTTTGGTCCCTCCCTTTGTTATGGAGCTCTGTTTCACTCTATTCTACTCT 57297
Qy 601 ATTAATCTTGGCACTGAAAAAATAAAAAA 632
Db 57296 ATTAATCTTGGCACTGAAAAAATAAAAAA 57265

RESULT 12
AL583805 89728 bp DNA linear PRI 29-SEP-2001
LOCUS Human DNA sequence from clone RP11-134K1 on chromosome 9, complete
DEFINITION sequence.
AC051658
AL583805 AC051658
VERSION AL583805.7 GI:15865009
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 89728)
Clark, G.
Direct Submission
JOURNAL Submitted (29-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Oct 2, 2001 this sequence version replaced gi:14702155.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-134K1 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-134K1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP11-134K1 is at 89728 in this
sequence. The true right end of clone RP11-60C15 is at 2000 in this
sequence.

```

```
FEATURES
  source      Location/Qualifiers
1. .89728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-134K1"
/clone_lib="RPC1-11.1"

ORIGIN
Query Match      88.9%; Score 564.6; DB 9; Length 89728;
Best Local Similarity 94.2%; Pred. No. 5.4e-167;
Matches 598; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTTTTAAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
|
|
|
Db 83434 CCTGTATCTTTAACTCTCTTTTAAAGTTTGTCTCTTCAGAAATCAAACTGTAA 83490
|
|
|
QY 61 CAAATTTGTTCTTCAATGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120
|
|
|
Db 83491 TATATGGTTCTTCAATGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGTGG 83550
|
|
|
QY 121 CTGACCGGCGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCACCCCTCCCGAG 180
|
|
|
Db 83551 CTGACCGGCGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCACCCCTCC 83610
|
|
|
QY 181 GAAATCTCAACTGCGCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACCGGT 240
|
|
|
Db 83611 GAAATCTCAACTGCGCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACCG 83670
|
|
|
QY 241 CATCAGCCAACTCTCCCAACAGCAGCTTGGGTTTCTCTGTGACAGGGGGGACTTGAGAGAC 300
|
|
|
Db 83671 TGTGAGCCAACTCTCCCAACAGCAGCTTGGGTTTCTCTGTGACAGGGGGGACTTGAGAG 83730
|
|
|
QY 301 AGGACTAGCTGGATTTCCTAGCCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
|
|
|
Db 83731 AGGACTAGCTGGATTTCCTAGCCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 83790
|
|
|
QY 361 GCATCCACCTCTAATCATGGGCTTGCACTTAGCTTCACCCGACCAATCAGAGAGCTC 420
|
|
|
Db 83791 GCATCCACCTCTAATCATGGGCTTGCACTTAGCTTCACCCGACCAATCAGAGAGCTC 83850
|
|
|
QY 421 ACTAAATCTTAATAGGCAGAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCCTG 480
|
|
|
Db 83851 ACTAAATCTTAATAGGCAGAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCCTG 83910
|
|
|
QY 481 AGAGCACAGCGGAGGACAAGATCGGATATAAACCCAGGCAATTCGAGCGCGCAACGG 540
|
|
|
Db 83911 AGAGCACAGCGGAGGACAAGATCGGATATAAACCCAGGCAATTCGAGCGCGCAACAG 83970
|
|
|
QY 541 CAACCCCTTTGGTCCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
|
|
|
Db 83971 CAACCCCTTTGGTCCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 84030
|
|
|
QY 601 ATTAATCTTGAACCTGAAAAAAGAAAAAAGAAAAA 635
|
|
|
Db 84031 ATTAATCTTGAACCTGAAAAAAGAAAAAAGAAAAA 84065
|
|
|

RESULT 13
AC072023
LOCUS
DEFINITION
Homo sapiens 3 BAC RP11-3J2 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
ACCESSION
AC072023
VERSION
AC072023.9 GI:27764634
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 112405)
Muzny,D.M., Adams,C., Ali-Osman,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,K.J.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homai,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,K., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulesed,H., Lozado,R., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
Sonaite,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vazquez,L.,
Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (07-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (22-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (30-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 112405)
Worley,K.C.
Direct Submission
Submitted (16-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jan 16, 2003 this sequence version replaced gi:21539113.
```

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

FEATURES

source	Location/Qualifiers
	1. .112405
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-3J2"
misc_feature	1. .2000
	/notes="overlaps bases 38072..40071 of clone AC125608"
	/function="clone overlap"
repeat_region	complement(1..261)
repeat_region	/rpt_family="L1M4"
repeat_region	complement(431..913)
repeat_region	/rpt_family="L1ME"
repeat_region	complement(914..1221)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(1222..1808)
repeat_region	/rpt_family="L1ME"
repeat_region	complement(2045..2198)
repeat_region	/rpt_family="L1ME"
repeat_region	2454..2811
repeat_region	/rpt_family="THE1C"
repeat_region	2813..3092
repeat_region	/rpt_family="MLT1I"
repeat_region	3194..3870
repeat_region	/rpt_family="L1PA3"
repeat_region	3871..3891
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(4297..4671)
repeat_region	/rpt_family="MSTD"
repeat_region	4714..4913
repeat_region	/rpt_family="MLT1I"
repeat_region	5632..5670
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(6337..6630)
repeat_region	/rpt_family="AluSx"

repeat_region	7063..7093	/rpt_family="TTTTA)n"
repeat_region	complement(7094..7374)	/rpt_family="AluSp"
repeat_region	complement(7396..7695)	/rpt_family="AluSx"
repeat_region	7910..8035	/rpt_family="MIR"
repeat_region	8824..10387	/rpt_family="L1PA16"
repeat_region	10388..10412	/rpt_family="AT_rich"
repeat_region	10637..10675	/rpt_family="AT_rich"
repeat_region	10754..10818	/rpt_family="L1PB1"
repeat_region	complement(10830..10905)	/rpt_family="L1PA16"
repeat_region	complement(10956..11184)	/rpt_family="L1"
repeat_region	11295..11658	/rpt_family="L2"
repeat_region	complement(12489..12812)	/rpt_family="L1PA15"
repeat_region	complement(12819..13533)	/rpt_family="L1"
repeat_region	13595..13619	/rpt_family="AT_rich"
repeat_region	14620..14933	/rpt_family="AluSg"
repeat_region	complement(15183..15254)	/rpt_family="L2"
repeat_region	complement(16463..17175)	/rpt_family="L1P"
repeat_region	17173..18635	/rpt_family="L1PA7"
repeat_region	18793..19462	/rpt_family="L1M4"
repeat_region	19461..20130	/rpt_family="L1ME"
repeat_region	20776..21073	/rpt_family="AluSx"

Query Match 88.2%; Score 560.2; DB 9; Length 112405;
Best Local Similarity 93.8%; Pred. No. 1.4e-165;
Matches 594; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY	1	CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTTCCAGAACTCAAACTGTAATACTA	60
DB	67581	CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTTCCAGAACTCAAACTGTAATACTA	67640
QY	61	CAAAATGTTCTTCAAAATGGAGACACAGATGGAGTCCATGACTAAGTCCACCGTGACCC	120
DB	67641	CAAAATGTTCTTCAAAATGGAGACACAGATGGAGTCCATGACTAAGTCCACCGTGACCC	67700
QY	121	CTGGACCGCTCTAGCCATGCTCCGATGTTTAAATGACATTAAGGACACCCCTCCGAG	180
DB	67701	CTGGACCGCTCTAGCCATGCTCCGATGTTTAAATGACATTAAGGACACCCCTCCGAG	67760
QY	181	GAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTGTAGACGGT	240
DB	67761	GAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTGTAGACGGT	67820
QY	241	CATCAGCCAACTCCCAACAGCACTTGGGTTTCTGTGTAGAGGGGGGAGTGGAGAC	300
DB	67821	CATCAGCCAACTCCCAACAGCACTTGGGTTTCTGTGTAGAGGGGGGAGTGGAGAC	67880
QY	301	AGGACTAGCTGGATTTCTTAGCCCAAGAGATCCCTAGCTAGCTAGCTAGCTAGCTAGCT	360
DB	67881	AGGACTAGCTGGATTTCTTAGCCCAAGAGATCCCTAGCTAGCTAGCTAGCTAGCTAGCT	67940
QY	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTACCTCAGCCCGGACCAATTCAGAGCTC	420


```
Db 67941 GCATCCACCTTTAARACACGGGCTCGCAACTTAGTCTCACACCAACCAATCAGAGAGCTC 68000
Qy 421 ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCTG 480
Db 68001 ACTAAATGCTTAATTAGGCAAAACAGAGGTAAA-TAATAGCCAATCATCTATTGCTG 68059
Qy 481 AGAGCACAGCGGAGGACAGAGTTCGGATATAAACCCAGGCATTCGAGCGGCAACGG 540
Db 68060 AGAGCACAGCGGAGGACAGAGTATATAAACAGGCATTCGAGCTGCGCATGG 68119
Qy 541 CAACCCCTTTGGGTCCTCCCTTTGATGGCGCTCTGTTTTCACCTCTATTTCACCT 600
Db 68120 CAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCACCTCTATTTCACCT 68179
Qy 601 ATTAATCTTGCACCTGAAAAAATAAAAAA 633
Db 68180 ATTAATCTTGCACCTGCAAAAAAATAAAAAA 68212

RESULT 14
BS000045/c
LOCUS
DEFINITION
BS000045 179114 bp DNA linear PRI 12-JUN-2004
Pan troglodytes chromosome 22 clone:RP43-179P23, map 22, complete
sequences.
ACCESSION BS000045 BA000046
VERSION BS000045.1 GI:37537312
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE
1 The International Chimpanzee Chromosome 22 Consortium.
DNA sequence and comparative analysis of chimpanzee chromosome 22
Nature 429, 382-388 (2004)
2 (bases 1 to 179114)
Hattori, M., Toyoda, A., Watanabe, H., Taylor, T. D., Kuroki, Y.,
Fujiyama, A. and Sakaki, Y.
Direct Submission
Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chinese National Human Genome Center at Shanghai, Shanghai, China;
*GFP, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
of Molecular Biotechnology, Jena, Germany; *KRIIB Genome Research
Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.

----- Genomic Center
Center: RIKEN Genomic Sciences Center
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: The Chimpanzee Chromosome 22 Sequencing Project
Center clone name: RP43-179P23
----- Summary Statistics
Sequencing vector: pUC18, pUC13, pTZ19; 100% of reads Chemistry:
Phrap; version 0.990329
Dye-terminator Big Dye and ET; 100% of reads Assembly program:
Consensus quality: 178,810 bases at least Q40
Consensus quality: 263 bases at least Q30
Consensus quality: 35 bases at least Q20
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
```

```
30);
an attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at one
plasmid
subclone or more than one M13 subclone;
and the assembly was confirmed by restriction digest.
-----
Source information:
The RPCI-43 chimpanzee BAC library was prepared from DNA isolated
from the blood of a single male chimpanzee using published
protocols (Osoegawa, K. et al. Genomics 52:1-8). The DNA from the
chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
characteristics are described at
http://www.chori.org/bacpac/mchimp43.htm.
The clone may be obtained from Pieter J. de Jong and coworkers
(http://www.chori.org/bacpac).
VECTOR: pBACe3.6
The CHORI-251 chimpanzee BAC library was prepared from DNA isolated
from the blood of a single male chimpanzee using published
protocols (Osoegawa, K. et al. Genomics 52:1-8). The DNA from the
chimpanzee ('Clint') was obtained from the Yerkes Primate Center in
Atlanta. The library was prepared by Baoli Zhu, Chung Li Shu,
Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library
characteristics are described at
http://www.chori.org/bacpac/chimpanzee251.htm.
The clone may be obtained from Pieter J. de Jong and coworkers
(http://www.chori.org/bacpac).
VECTOR: pTARBAC2.1
The PFB1 chimpanzee BAC library was prepared from DNA isolated from
cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKSI145
The PTF22 chimpanzee Fosmid library was prepared from DNA isolated
from cultured cells established from the blood of a single male
chimpanzee.
Clones may be obtained from Asao Fujiyama and co-workers
(http://www.gsc.riken.go.jp).
VECTOR: pKSI143
-----
Sequence Quality Assessment:
This entry has been annotated with sequence
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in
10,000 bp.
-----
Neighboring clones: RP43-014B20(left) and PTB-042H12(right).
-----
FEATURES
source
1. 179114
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="22"
/clone="RP43-179P23"
/clone_lib="RPCI-43 chimpanzee BAC"
-----
ORIGIN
Query Match 88.2%; Score 559.8; DB 9; Length 179114;
Best Local Similarity 94.0%; Pred. No. 2e-165;
Matches 582; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Qy 1 CCCGTATCTTTAACTCCCTTGTAAAGTTGCTCTCTCCAGAAATCAAACTGTAACACTA 60
Db 169185 CCGTGTATCTTTAACTCCCTTGTAAAGTTGCTCTCTCCAGAAATCAAACTGTAACACTA 169126
Qy 61 CAAATTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGAGCC 120
Db 169125 CAAATAGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGAGCC 169066
```


QY 121 CTGACCGGCTGTAGCCCTGCTCCGATGTTATGACATTTGAAGGACACCCCTCCCGAG 180
 DB 169065 CTAGACTGGCTGTAGCCCTGCTCCGATGTTATGACATTTGAAGGACACCCCTCCCGAG 169006
 QY 181 GAAATCTCAACTGCAACAACCCCTACTATGCCCCAATTCAGCGGGAGCAGTTAGAGCGGT 240
 DB 169005 GAAATCTCAACTGCAACAACCCCTACTATGCCCCAATTCAGCGGGAGCAGTTAGAGCGGT 168946
 QY 241 CATGAGCAACCTCCCGACAGCACTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 300
 DB 168945 CATGAGCAACCTCCCGACAGCACTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 168886
 QY 301 AGGACTAGCTGATTTCTTAGGCCCAACGAAGATCCCTAAGCTAGCTGGGAAGTGTACT 360
 DB 168895 AGGACTAGCTGATTTCTTAGGCCCAACGAAGATCCCTAAGCTAGCTGGGAAGTGTACT 168826
 QY 361 GCATCCACTCTTAACATGGGGCTTGGCACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
 DB 168825 GCATCCACTCTTAACATGGGGCTTGGCACTTAGCTCAACCCGACCAATCAGAGAGCTC 168766
 QY 421 ACTAAATGCTTAATTTAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATTTGCGCTG 480
 DB 168765 ACTAAATGCTTAATTTAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATTTGCGCTG 168706
 QY 481 AGAGCACAGCGGAGGACCAAGATCGGGATATAAACCCAGGCACTTCAGAGCGGCAACGG 540
 DB 168705 AGAGCACAGCGGAGGACCAAGATCGGGATATAAACCCAGGCACTTCAGAGCGGCAACGG 168646
 QY 541 CAACCCCTTTGGGCTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTATTTCACTCT 600
 DB 168645 CAACCCCTTTGGGCTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTATTTCACTCT 168586
 QY 601 ATTAATCTTGAACCTGAA 619
 DB 168585 ACTAAATCTTGAACCTGCA 168567

RESULT 15
 LOCUS AP001538 174019 bp DNA linear PRI 25-MAR-2000
 DEFINITION Homo sapiens genomic DNA, chromosome 21q21.1-q21.2 clone:B853K11,
 LL56-APP region, complete sequence.
 ACCESSION AP001538
 VERSION AP001538.1 GI:7328982
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 174019)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Watanabe,H., Yada,T.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
 Homo sapiens 174,019 genomic DNA of 21q21.1-q21.2
 REFERENCE 2 (bases 1 to 174019)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
 Direct Submission
 Submitted (23-MAR-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail:hattori@gsc.riken.go.jp)
 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
 Fax:81-42-778-9924

FEATURES
 source
 Location/Qualifiers
 1. 174019
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q21.1-q21.2"
 /clone="B853K11"

ORIGIN

Query Match 87.7%; Score 556.6; DB 9; Length 174019;
 Best Local Similarity 93.7%; Pred. No. 2.1e-164;
 Matches 580; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 1 CCCTGTATCTTTAAACCTCCCTTGTAAAGTTTCTCTCCAGAAATCAAAACCTGTAATAACTA 60
 DB 36710 CCCTGTATCTTTAAACCTCCCTTGTAAAGTTTCTCTCCAGAAATCAAAACCTGTAATAACTA 36651
 QY 61 CAAATGTTCTTCAAAATGAGACACAGATGAGATCCATGATTAAGATCCACCGTGGACCC 120
 DB 36650 CAAATGTTCTTCAAAATGAGACACAGATGAGATCCATGATTAAGATCCACCGTGGACCC 36591
 QY 121 CTGACCGGCTGTAGCCCAATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 180
 DB 36590 CTGACCGGCTGTAGCCCAATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 36531
 QY 181 GAAATCTCAACTGCAACAACCCCTACTATGCCCCAATTCAGCGGGAGCAGTTAGAGCGGT 240
 DB 36530 GAAATCTCAACTGCAACAACCCCTACTATGCCCCAATTCAGCGGGAGCAGTTAGAGCGGT 36471
 QY 241 CATCAGCAACCTCCCGACAGCACTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 300
 DB 36470 CATCAGCAACCTCCCGACAGCACTGGGTTTCTGTTGAGGGGGGACTGAGAGAC 36411
 QY 301 AGGACTAGCTGATTTCTTAGGCCCAACGAAGATCCCTAAGCTAGCTGGGAAGTGTACT 360
 DB 36410 AGGACTAGCTGATTTCTTAGGCCCAACGAAGATCCCTAAGCTAGCTGGGAAGTGTACT 36351
 QY 361 GCATCCACTCTTAACATGGGGCTTGGCACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
 DB 36350 GCATCCACTCTTAACATGGGGCTTGGCACTTAGCTCAACCCGACCAATCAGAGAGCTC 36291
 QY 421 ACTAAATGCTTAATTTAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATTTGCGCTG 480
 DB 36290 ACTAAATGCTTAATTTAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATTTGCGCTG 36231
 QY 481 AGAGCACAGCGGAGGACCAAGATCGGGATATAAACCCAGGCACTTCAGAGCGGCAACGG 540
 DB 36230 AGAGCACAGCGGAGGACCAAGATCGGGATATAAACCCAGGCACTTCAGAGCGGCAACGG 36171
 QY 541 CAACCCCTTTGGGCTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTATTTCACTCT 600
 DB 36170 CAACCCCTTTGGGCTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTATTTCACTCT 36111
 QY 601 ATTAATCTTGAACCTGAA 619
 DB 36110 ATTAATCTTGAACCTGCA 36092

RESULT 16
 LOCUS AP001674/c 340000 bp DNA linear PRI 21-MAY-2003
 DEFINITION Homo sapiens genomic DNA, chromosome 21q, section 18/105.
 ACCESSION AP001674 AL163219 BA000005
 VERSION AP001674.1 GI:7768666
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
 Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
 Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
 Polley,A., Menzel,U., Delabar,J., Rump,A., Schillhabel,M., Schudy,A.,
 Patterson,D., Reichwald,K., Kudoh,J., Shibuya,K., Kawasaki,K.,
 Zimmermann,W., Rosenthal,A., Kudo,H., Nagamine,K., Mitsuayama,S.,
 Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Nordsiek,G.,
 Antonarakis,S.E., Minoshima,S., Shimizu,N., Desario,A.,
 Hornischer,K., Barandt,P., Scharfe,M., Schoen,O., Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
 Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,

Fri Feb 25 16:26:29 2005

us-09-319-156b-6.rge

Gardiner, K., Nizetic, D., Francis, F., Leirach, H., Reinhardt, R. and Yaspo, M.L.
The DNA sequence of human chromosome 21

Nature 405 (6784), 311-319 (2000)

20289799

10830953

2 (bases 1 to 340000)

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Hattori, M., Fujiyama, A., Taylor, T.D., Watanabe, H., Yada, T., Park, H.S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D.K., Soeda, E., Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., Blechschmidt, K., Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R., Patterson, D., Reichwald, K., Rump, A., Schillhabel, M., Schudy, A., Zimmermann, W., Rosenthal, A., Kudoh, J., Nagamine, K., Mitsuyama, S., Asakawa, S., Shintani, A., Sasaki, T., Shimizu, N., Nordstiek, G., Antonarakis, S.E., Minoshima, S., Scharfe, M., Schoen, O., Desario, A., Hornischer, K., Barand, P., Bloeker, H., Ramser, J., Beck, A., Klages, S., Reichelt, J., Kauer, G., Bloeker, H., Ramser, J., Beck, A., Klages, S., Hennig, S., Riesselmann, L., Dagand, E., Wehrmaier, S., Borzym, K., Gardiner, K., Nizetic, D., Francis, F., Leirach, H., Reinhardt, R. and Yaspo, M.L.

Direct Submission

Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing Consortium: * RIKEN Genomic Sciences Center, Human Genome Research Group * Institute of Molecular Biotechnology, Genome Analysis * Keio University School of Medicine, Dept. of Molecular Biology * GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular Genetics (addresses see below)

On May 30, 2000 this sequence version replaced gi:7717271.

The chromosome 21 mapping and sequencing consortium consisting of

* RIKEN Genomic Sciences Center, Human Genome Research Group, * Sagamihara 228-8555, Japan,

* e.mail: hattori@gsc.riken.go.jp

* URL: http://hgp.gsc.riken.go.jp/

and

* Institute of Molecular Biotechnology, Genome Analysis, * Beutenbergstrasse 11, D-07745 Jena, Germany,

* e.mail: gscj-submit@genome.imb-jena.de

* URL: http://genome.imb-jena.de/

and

* Keio University School of Medicine, Molecular Biology, * Tokyo 160-8582, Japan,

* e.mail: nehimizu@mb-med.keio.ac.jp

* URL: http://www.dmb.med.keio.ac.jp/

and

* GBF, Dept. of Genome Analysis,

* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail: info.genome@gbf.de

* URL: http://genome.gbf.de/

and

* Max-Planck Institute for Molecular Genetics,

* Innestrasse 73, D-14195 Berlin, Germany,

* e.mail: info-chr21@molgen.mpg.de

* URL: http://chr21.rz-berlin.mpg.de/

AL163219: Submitted (10-Apr-2000).

FEATURES

source

1..340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

<1..6149

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R749F18, 5' partial"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP001537"

<1..165231

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R853K11, 5' partial"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP001538"

136796..315056

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R845I12"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AF245226"

156433..277809

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="P646P19"

/clone_lib="RPCI1.3-5 PAC library"

/note="Accession No. AL157359"

207594..250588

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="Q3D19"

/clone_lib="LL21NCO3-Q Cosmid library"

/note="Accession No. AL049753"

276958..301481

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R755I19"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP001345"

299634..>340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R28M9, 3' partial"

/clone_lib="RPCI-11 BAC library"

/note="Accession No. AP000431"

311751..>340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="C103E2331, 3' partial"

/clone_lib="ICRFcl03 Cosmid library"

/note="Accession No. AL132993"

15..444

/note="L1MA4A"

/rpt_family="LINE/L1"

/rpt_type="DISPERSED"

555..3531

/note="L1MA4A"

/rpt_family="LINE/L1"

/rpt_type="DISPERSED"

3770..3876

/note="L2"

/rpt_family="LINE/L2"

/rpt_type="DISPERSED"

complement(5645..6178)

/note="L1MA2"

repeat_region

repeat_region

repeat_region

repeat_region

```
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
6181..6534
/notes="THE18"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
8021..8370
/notes="THE12"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(8373..10032)
/notes="LiMA2"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
10213..10520
/notes="AluJb"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
15810..15851
/notes="(CA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(16628..16697)
/notes="MER53"
/rpt_family="DNA"
/rpt_type=DISPERSED
16638..16715
/notes="MER53"
/rpt_family="DNA"
/rpt_type=DISPERSED
18847..19585
/notes="L2"
/rpt_family="LINE/L2"

Query Match      87.7%; Score 556.6; DB 9; Length 340000;
Best Local Similarity 93.7%; Pred. No. 2.3e-164;
Matches 580; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCCTCTGTTAGTTTGTTCTCTCCAGATCAAACTGTAAACTA 60
DB 27921 CCCTGTATCTTTAACTCCTCTGTTAGTTTGTTCTCTCCAGATCAAACTGTAAACTA 27862

QY 61 CAAATTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCCGTGGACCC 120
DB 27861 CAAATGTTCTTCAATGGGGCCCCAGATGCACTGATGATCCACCCGGGCCC 27802

QY 121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 27801 CTAGACTGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCCAAG 27742

QY 181 GAAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 27741 GAAATCAACTGCCCCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCTGT 27682

QY 241 CATCAGCCACCTCCCAACAGCAGCTGGGTTTTCTGTTGAGGGGGGACTGAGAGAC 300
DB 27681 CATGGCCACACCTCCCAACAGCAGCTGGGTTTTCTGTTGAGGGGGGACTGAGAGAC 27622

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACT 360
DB 27621 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACC 27562

QY 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGCTC 420
DB 27561 GCATCCACCTTTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGCTC 27502

QY 421 ACTAAATGCTAATTAGGCAAAATAGAGGTAAAGAAATAGCAATCATCTATTGGCTG 480
DB 27501 ACTAAATGCTAATTAGGCAAAATAGAGGTAAAGAAATAGCAATCATCTATTGGCTG 27442

QY 481 AGAGCAGCGGGAGGAGCAGGATCGGGATATAAACCAGGATTCGAGCCGGCAGACGG 540
DB 27441 AGAGCAGCAGGAGGAGCAGGATCAGGATGTAACCCAGGCTTCAAGCTGCGACGG 27382
```

```
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATTCTACTCT 600
DB 27381 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATTCTACTCT 27322

QY 601 ATTAAATCTTGCACACTGAA 619
DB 27321 ACTAAATCTTGCACACTGCA 27303

RESULT 17
AP001545 149755 bp DNA linear HTG 13-JUL-2000
Homo sapiens chromosome 18 clone RP11-762G24 map 18q21, WORKING
DRAFT SEQUENCE, 13 unordered pieces.
AP001545
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 149755)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
Homo sapiens 149,755 genomic DNA of 18q21
Published Only in Database (2000)
2 (bases 1 to 149755)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.,
Direct Submission
Submitted (24-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
On Jul 14, 2000 this sequence version replaced gi:8117383.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: RP11-762G24
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 145596 bases at least Q40
Consensus quality: 147102 bases at least Q30
Consensus quality: 147971 bases at least Q20
Insert size: 148555; sum-of-contigs
Quality coverage: 10.98x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
13 contigs. The true order of the pieces is not known and the
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 44828 contig of 44828 bp in length
44929 82602 contig of 37674 bp in length
82703 101344 contig of 18642 bp in length
101445 116205 contig of 14761 bp in length
116306 123784 contig of 7479 bp in length
123885 130355 contig of 6471 bp in length
130456 133619 contig of 3164 bp in length
133720 137416 contig of 3697 bp in length
137517 141540 contig of 4024 bp in length
141641 144351 contig of 2711 bp in length
144542 146667 contig of 2216 bp in length
```

	Matches	583;	Conservative	0;	Mismatches	52;	Indels	0;	Seqs
QY	1	CCCTGTATCTTTAACTCCTTGTTAGTGTCTCTTCCAGAATCAAACTGTAATACTA	60						
Db	67067	CCCTGTATCTTTAACTCCTTGTTAGTGTCTCTTCCAGAATCAAAAGCTGTAAATACTA	67126						
QY	61	CAAAATTGTTCTTCAAAATGGAGCACACAGATGGATCTCATGACTAAGATCCACCGTGGACCC	120						
Db	67127	CAAAATCATCTTCAAAATGGAGCCCCAGATGCAGTCCCATGACTAAGATCTTACCACAGACCC	67186						
	131	CTGGACACCGGCTGTAGTCCCATGCTCCGATCTTAAATGACATTTGAAGGCACCCCTCCCGAG	180						

121	QY	CTGGACCGCGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGACACCTCTCCCAAG	180
67187	Db	CTGGACACGCGCTGCTAGCCCATGCTCCGATGTTTAATGACATCGAAGGCAACCCCTCCCAAG	67246
181	QY	GAATCTCAACTGTCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTTACAGCGGT	240
67247	Db	GAATCTCAACTGTCACAAACCCCTACTATGCCCCAAGTTACAGCAGGAAGCAGTTTAGGCGGT	67306
241	QY	CATCAGCCAACTCCCCCAACAGCACTTGGGTTTTCTGTTCAGAGGGGGGACTCAGAGAC	300
67307	Db	CATCAGCCAACTCCCCCAACAGCACTTGGGTTTTCTGTTCAGAGGGGGTACTCAGAGAC	67366
301	QY	AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT	360
67367	Db	AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAACCTAGCTGGGAAGGTGACT	67426
361	QY	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420
67427	Db	GCATTCACCTTTAAACACGAGGCTTGCACCTTAGCTCACAGCTGACCAATCAGAGAGCTC	67486
421	QY	ACTAAATGCTAATTAGGCCAAAAATAGAGGGTAAAGAAATAGCCCAATCATCTATTGCGCTG	480
67487	Db	ACTAAATGCTAATTAGGCCAAAAACAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG	67546
481	QY	AGAGCACAGGGGAGGACAAGGATCGGATATAAAACCCAGGCATTCGAGCGCGCAACGG	540

67547 AGAGCACGTGGGAGGACAGGATCAGGATATAACCCAGGCATTCGAGCCGGCAACGG 67606

DB	67547	AGAGCACAATGGGAGCGACCTCCCTTTGTATGGGCCTCTGTTTTCACTCTATTTCACCTCT	600
Qy	541	CAACCCCCTTGGTGCCCTCCCTTTGTATGGGCCTCTGTTTTCACTCTATTTCACCTCT	
DB	67607	CAACCCCCTTGGTGCCCTCCCTTTGTATGGGAGCTCTGTTTTCACTCTATAAATCTT	67666
Qy	601	ATTAAATCTTGCACCTGAATAAAAAAAAAAAAAA	635

Db	67667	TCAACTGC	AAAAAAAAAT	AAAAAAAAAT	AAAAAAAAAT	AAAAA	67701
RESULT 18	AC090313/c						
LOCUS	AC090313	152980 bp	DNA	linear	HTG 27-MAR-2003		
DEFINITION	Homo sapiens chromosome 18 clone RP11-762G24	map 18.					
ACCESSION	AC090313						
VERSION	AC090313.12	GI:2123575					
KEYWORDS	HTG; HTGS PHASE2; HTGS_FULLTOP; HTGS_CANCELLED.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	1 (bases 1 to 152980)						
TITLE	Biren,B., Nusbaum,C. and Lander,E.						
JOURNAL	Homo sapiens chromosome 18, clone RP11-762G24						
REFERENCE	Unpublished						
AUTHORS	2 (bases 1 to 152980)						
	Biren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,						
	Barna,N., Bastien,V., Boguslavsky,L., Bouckhalter,B., Brown,A.,						
	Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,						
	Collumore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,						
	Dodgess,S., Farcio,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,						
	Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,						
	Harada,S., Haxel,S., Horton,L., Hulme,W., Iliev,I., Johnson,R.,						
	Johnson,R., Johnson,R., Johnson,R., Johnson,R., Johnson,R.,						

Hagos, B., Heatold, A., Horton, E., Hume, J., Jones C., Karatas, A., LaRocque, K., Lamazares, R., Landers, T.,

Lehoczky, J., Levine, K., Liu, G., Maclean, P., MacKinnon, S., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,

McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, J., Navlor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strausz, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barn, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgaiter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Navlor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE
Direct Submission
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:21699527.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12582
Center clone name: 762_G_24

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 152980: contig of 152980 bp in length.

FEATURES
source

1. 152980
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-762G24"
/clone_lib="RP11 Human Male BAC"

ORIGIN

Query Match 86.9%; Score 551.8; DB 2; Length 152980;
Best Local Similarity 91.8%; Pred. No. 6.8e-163;
Matches 583; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTTAACTCTCTTGTGTTTAAAGTTTGTCTTCTCAGAAATCAAAACCTGTAAACCTA 60
Db 48688 CCCTGTATCTTTTAACTCTCTTGTGTTTAAAGTTTGTCTTCTCAGAAATCAAAACCTGTAAACCTA 48629
QY 61 CAAATGTTCTTCAATGGAGACACAGATGGAGTCCATGACTTAAGTCCACCGTGGACCC 120
Db 48628 CAAATCATTTCTTCAATGGAGACACAGATGGAGTCCATGACTTAAGTCCACCGTGGACCC 48569
QY 121 CTGGACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTTGAAGGACACCCCTCCCGAG 180
Db 48569 CTGGACCGGCTGTAGCCATGCTCCGATGTTTAATGACATTTGAAGGACACCCCTCCCGAG 48509
QY 181 GAAATCTCAACTGSCAACACCCCTACTATGCCCCAATTCAGCGGAGAGCTAGAGCGGT 240
Db 48508 GAAATCTCAACTGSCAACACCCCTACTATGCCCCAATTCAGCGGAGAGCTAGAGCGGT 48449
QY 241 CATCAGCCAACTCTCCCAACAGACTTGGTTTCTGTTGTAGAGGGGGGAGCTGAGAGAC 300
Db 48448 CATCAGCCAACTCTCCCAACAGACTTGGTTTCTGTTGTAGAGGGGGGAGCTGAGAGAC 48389
QY 301 AGGACTAGCTGATTTCTTAGGCCCAAGAAATCCCTAAGCTTAGCTGGGAGGTGACT 360
Db 48388 AGGACTAGCTGATTTCTTAGGCCCTGACTGACTGACTGACTGACTGACTGACTGACT 48329
QY 361 GCATCCACCTCTAAACATGGGGCTTGCACCTTAGCTCACCACCGACCAATCAGAGAGCTC 420
Db 48328 GCATCCACCTCTAAACATGGGGCTTGCACCTTAGCTCACCACCGACCAATCAGAGAGCTC 48269
QY 421 ACTAAATGCTAATTTAGGCAAAATAGGAGTTAAGAAATAGCAATCATCTATTTCCTCTG 480
Db 48268 ACTAAATGCTAATTTAGGCAAAATAGGAGTTAAGAAATAGCAATCATCTATTTCCTCTG 48209
QY 481 AGAGCACACGGGAGGACCAAGATCGGATATAAACCCAGGATTCAGAGCCGCGCAACGG 540
Db 48208 AGAGCACACGGGAGGACCAAGATCGGATATAAACCCAGGATTCAGAGCCGCGCAACGG 48149
QY 541 CAACCCCTTTTGGGCTCCCTCTCTTGTATGGCGCTCTGTTTCACTCTATTTCCTCTCT 600
Db 48148 CAACCCCTTTTGGGCTCCCTCTCTTGTATGGCGCTCTGTTTCACTCTATTTCCTCTCT 48089
QY 601 ATTAATCTTGCACCTGAAAAAATAAAAAATAAAAAATAAAAA 635
Db 48088 TCAACTGCAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 48054

RESULT 19
AC093531/c
LOCUS
DEFINITION Homo sapiens chromosome 5 clone RP11-405L7, complete sequence.
ACCESSION AC093531
VERSION AC093531.2 GI:16945981
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163803)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163803)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 163803)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Nov 16, 2001 this sequence version replaced gi:15383820.
Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.3.
Location/Qualifiers
1. .163803
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-405L7"

FEATURES
source

ORIGIN

Query Match 86.9%; Score 551.8; DB 9; Length 163803;
Best Local Similarity 93.2%; Pred. No. 6.9e-163;
Matches 577; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
Db CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTTCAGAAATGAAAGCTGTAAAGCTA 117289

QY 61 CAATTTGTTCTTCAATGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db CAATAGTTCCTCAATGGAACCCAGATGCGAGTCCATGACTAATCTACCGTGGACCC 117229

QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 180
Db CTGGACCGGCTGCTAGATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAG 117169

QY 181 GAAATCTCAACTGCACAACCCCTAATATCCCCCAATTCAGCGGAGCAGTTAGACCGGT 240
Db GAAATCTCAACTGCACAACCCCTACTACACTCCATTCAGTAGGAGCAGTTAGACGAGT 117109

QY 241 CATCAGCCCACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGCACTGAGAGAC 300
Db TGTCAAGCCCACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGTGCATGAGAGAC 117049

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAAGCTAGCTGGGAAGTGACT 360
Db AGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCTAAGCTAGCTGGGAAGTGACC 116989

QY 361 GCATCACTCTTAAACATGGGCTTCACTTAGCTCACACCCGACCAATCAGAGGCTC 420
Db GCATCACTCTTAAACATGGGCTTGCACCTTAGCTCACACCCGACCAATCAGAGGCTC 116929

QY 421 ACTAAATGCTAATTAGGCACAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCCTG 480
Db ACTAAATGCTAATCAGGCACAAACAGGAGGTAAGCAATAGCCAAATCATCTATTGCCTG 116869

QY 481 AGAGCAGCGGAGGAGCAGGATCGGATATAAACCCAGGCAATTCGAGCCGGCAACGG 540
Db AGAGCAGCGGAGGAGCAGGATCGGATATAAACCCAGGCAATTCGAGCCAGCAACAG 116809

QY 541 CAACCCCTTTGGTCCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db CAACCCCTTTGGTCCCTCCCTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 116749

QY 601 ATTAATCTTGCACACTGAA 619
Db ATTAATCTTGCACACTGCA 116730

RESULT 20
AC107075/c
LOCUS AC107075 176425 bp DNA linear PRI 12-JUN-2002
DEFINITION Homo sapiens BAC clone RP11-452N17 from 2, complete sequence.
ACCESSION AC107075
VERSION AC107075.4 GI:21263352
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 176425)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 176425)
Swearengen-Shahid,S., Meyer,R. and Dignan,G.
The sequence of Homo sapiens BAC clone RP11-452N17
Unpublished (2001)
3 (bases 1 to 176425)
Waterston,R.H.
Direct Submission
Submitted (15-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 176425)
Waterston,R.H.
Direct Submission
Submitted (30-MAY-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 176425)
Waterston,R.
Direct Submission
Submitted (12-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 30, 2002 this sequence version replaced gi:18640713.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_NH0452N17

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-624P4; the clone sequenced
to the right is RP11-362J3. Actual start of this clone is at base
position 1 of RP11-452N17; actual end is at base position 176425 of
RP11-452N17.

Polymorphisms exist between AC110086 and AC107075. Data from AC110086, AC116626 and AC018872 was used to finish AC107075.

FEATURES

Source

1. 176425

/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"

/clone="RP11-452N17"
/clone_lib="RPCI-11"
2423. 2523
/rpt_family="MIR"

3564. 3598
/rpt_family="AT_rich"

4522. 4567
/rpt_family="AT_rich"

4679. 4756
/rpt_family="(TATAA)n"

5700. 5733
/rpt_family="(A)n"

5973. 6006
/rpt_family="(A)n"

6123. 6144
/rpt_family="AT_rich"

6803. 6896
/rpt_family="L1"

6923. 6999
/rpt_family="L1"

7230. 7715
/rpt_family="MaLR"

7852. 7966
/rpt_family="Tc2"

7966. 8026
/rpt_family="Tc2"

8043. 8172
/rpt_family="Tc2"

8174. 8247
/rpt_family="(TATATG)n"

8250. 8456
/rpt_family="Tc2"

8922. 9642
/rpt_family="L1"

11590. 11632
/rpt_family="AT_rich"

12173. 12321
/rpt_family="MER1_type"

12323. 12343
/rpt_family="(A)n"

13107. 14677
/rpt_family="L1"

14678. 14704
/rpt_family="(TTTG)n"

14708. 14996
/rpt_family="Alu"

14997. 15173
/rpt_family="L1"

15615. 15720
/rpt_family="CT-rich"

15811. 16031
/rpt_family="MaLR"

16093. 16330
/rpt_family="L1"

16724. 17718
/rpt_family="L1"

18878. 19038
/rpt_family="Alu"

19298. 19551
/rpt_family="Alu"

19552. 19912
/rpt_family="MaLR"

23674. 23705
/rpt_family="(CAGA)n"

repeat_region 25248. 25839
/rpt_family="MaLR"
repeat_region 27499. 27532
/rpt_family="AT_rich"
repeat_region 27693. 27840
/rpt_family="MER1_type"
repeat_region 27902. 28090
/rpt_family="MIR"
repeat_region 28091. 28254
/rpt_family="MER1_type"
repeat_region 28283. 28477
/rpt_family="MER1_type"
repeat_region 28510. 28533
/rpt_family="AT_rich"
repeat_region 28629. 28649
/rpt_family="AT_rich"
repeat_region 29707. 29751
/rpt_family="(TA)n"
repeat_region 30414. 30847
/rpt_family="L1"
repeat_region 31385. 31942
/rpt_family="L1"
repeat_region 32421. 32591
/rpt_family="L2"
repeat_region 32592. 32983
/rpt_family="MaLR"
repeat_region 32984. 33714
/rpt_family="L2"
repeat_region 34913. 35339
/rpt_family="MaLR"
repeat_region 35752. 36049
/rpt_family="Alu"
repeat_region 36274. 36330
/rpt_family="Alu"
repeat_region 36767. 36809
/rpt_family="AT_rich"
repeat_region 37396. 37437
/rpt_family="AT_rich"
repeat_region 37816. 37850
/rpt_family="A-rich"
repeat_region 38954. 39375
/rpt_family="L1"
repeat_region 40130. 40216
/rpt_family="L2"
repeat_region 41723. 42015
/rpt_family="Alu"

Query Match 86.9%; Score 551.8; DB 9; Length 176425;
Best Local Similarity 92.8%; Pred. NO. 6.9e-163;
Matches 593; Conservative 0; Mismatches 37; Indels 9; Gaps 1;

QY	1	CCCTGTATCTTTAACTCCTTGTGTTAGTTTGTCTTCCAGATCAAACTGTAACCTA	60
Db	161068	CCCTATATCTTTAACTCCTTGTGTTAGTTTGTCTTCCAGATCAAACTGTAACCTA	161009
QY	61	CAAAATGTTCTTCAAAATGGAGCAGATGGAGTCCATGATCAAGATCCACCGTGACCC	120
Db	161008	CAAAATGTTCTTCAAAATGGAGCAGATGGAGTCCATGATCAAGATCCACCGTGACCC	160949
QY	121	CTGGACCGGCTGTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCGGAG	180
Db	160948	CTGGACCGGCTGTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCAAG	160889
QY	181	GAATCTCAACTGCACAAACCCCTACTATGCCCAATTTACGCGGAGCAGCTTAGACGGT	240
Db	160888	GAATCTCAACTGCACAAACCCCTACTATGCCCAATTTACGCGGAGCAGCTTAGACGGT	160829
QY	241	CATCAGCAACCTCCCAACAGACATTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC	300
Db	160828	CTTCGCGCAACCTCCCAACAGACATTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC	160769
QY	301	AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAGCCCTAGCTGGGAAGGTGACT	360

Fri Feb 25 16:26:29 2005

us-09-319-156b-6.rge

Db 160768 AGGACTAGCTGATTTCTTAGGCCGCACTAAGATCCCTAAGCCTAGCTAGTGGGAAGTGACC 160709
Qy 361 GCATCCACTCTTAACATAGGGCTTGCACCTTAGCTCTACACCGCACCAATCA----- 412
Db 160708 GTTTCATCTTTAAACACGGGGCTTACAACTTAACCTCACATGACCAATCAGATAGTAA 160649
Qy 413 -GAGAGCTCACTAAATGCTAATTAGGCAGAAAATAGGAGGTAAAGAAATAGCCATCATC 471
Db 160648 GGAGAGCTCACTAAATGCTAATTAGGCAGCAACACAGGAGGTAAAGAAATAGCCATCACC 160589
Qy 472 TATTCCTGAGACGACGCGGAGGAGGACAAAGGATCGGATATAAACCCAGGCAATTCGAGC 531
Db 160589 TGTTCCTGAGACGACGCGGAGGAGGACATGATCGGATATAAACCCAGGCAATTCGAGC 160529
Qy 532 CGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTTA 591
Db 160528 CGGCAACGACAAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTGCTTTTCACTCTTA 160469
Qy 592 TTTCACTCTAATTAATCTTGCACCTGCAATGAAAAAARAAAA 630
Db 160468 CTTCACTCTATCAATCTTTGCAACTGCAAAAAAARAAAA 160430

RESULT 21
AL139090/c
LOCUS
DEFINITION
Human DNA sequence from clone RP11-12B13 on chromosome 6 Contains
STSs and GSSs, complete sequence.
ACCESSION
AL139090
VERSION
AL139090.11 GI:11228536
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 128468)
Garner, P.
Direct Submission
Submitted (05-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 20, 2000 this sequence version replaced gi:1125403.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/c-elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
IMPORTANT: This sequence is not the entire insert of clone
RP11-12B13 it may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP11-12B13 is at 1 in this sequence. The
true left end of clone RP1-125N20 is at 128369 in this sequence.
The true right end of clone RP11-13D18 is at 88802 in this
sequence. This sequence was finished as follows unless otherwise
noted: all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
RP11-12B13 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm

FEATURES
source
VECTOR: pBACE3.6.
Location/Qualifiers
1..128468
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-12B13"
/clone_lib="RPCI-11.1"
1..353
/note="match: GSS: Em:AQ609536"
7..448
/note="match: GSS: Em:AQ704228"
68..350
/note="match: GSS: Em:B75711"
571..1535
/note="LTRS repeat: matches 1..966 of consensus"
1966..2087
/note="L2 repeat: matches 2574..2704 of consensus"
2290..2509
/note="L2 repeat: matches 2098..2328 of consensus"
2976..3352
/note="L2 repeat: matches 1693..2089 of consensus"
3497..3668
/note="MER20 repeat: matches 16..188 of consensus"
complement(4521..5085)
/note="match: GSS: Em:AQ421033"
complement(4571..5086)
/note="match: GSS: Em:AQ883615"
complement(4633..5083)
/note="match: GSS: Em:AQ812769"
5012..5530
/note="MER41B repeat: matches 114..601 of consensus"
5081..5634
/note="match: GSS: Em:AQ544837"
6085..6174
/note="MIR repeat: matches 64..153 of consensus"
6235..6416
/note="MER5A repeat: matches 3..186 of consensus"
7146..8805
/note="L1ME3 repeat: matches 4381..6140 of consensus"
8806..9399
/note="L1PA5 repeat: matches 5526..6143 of consensus"
9400..9710
/note="L1ME3 repeat: matches 4094..4381 of consensus"
9711..15846
/note="L1PA6 repeat: matches 11..6143 of consensus"
15826..16269
/note="L1M4 repeat: matches 3692..4135 of consensus"
17077..17544
/note="L1M4 repeat: matches 2710..3168 of consensus"
17552..17740
/note="L1MEC repeat: matches 2408..2260 of consensus"
17741..18039
/note="AluXs repeat: matches 1..299 of consensus"
18040..18346
/note="L1MEC repeat: matches 2106..2409 of consensus"
18447..18874
/note="L1MEC repeat: matches 1469..1910 of consensus"
complement(18475..18995)
/note="match: GSS: Em:AQ755181"
18938..19072
/note="L1MEC repeat: matches 1243..1377 of consensus"
19459..19514
/note="28 copies 2 mer ca 96% conserved"
20321..20675
/note="LTR16B repeat: matches 97..464 of consensus"
20952..21137
/note="AluJo repeat: matches 120..307 of consensus"
21203..21254
/note="26 copies 2 mer aa 71% conserved"
23289..23682
/note="match: GSS: Em:AQ881909"

repeat_region 23421..23472
 misc_feature /note="MIR repeat: matches 90..141 of consensus"
 repeat_region 24395..24791
 repeat_region /note="match: GSS: Em:AQ027280"
 repeat_region 24638..24848
 repeat_region /note="HAL1 repeat: matches 428..638 of consensus"
 repeat_region 25008..25217
 repeat_region /note="105 copies 2 mer aa 55% conserved"
 repeat_region 23328..25391
 repeat_region /note="32 copies 2 mer ta 70% conserved"
 repeat_region 25482..26261
 repeat_region /note="LIP3 repeat: matches 5..776 of consensus"
 repeat_region 26257..31504
 repeat_region /note="LIP3 repeat: matches 900..6146 of consensus"
 repeat_region 31614..32312
 repeat_region /note="LIM4 repeat: matches 3918..4608 of consensus"
 repeat_region 32398..32650
 repeat_region /note="LIM4 repeat: matches 4626..4909 of consensus"
 repeat_region 32799..32963
 repeat_region /note="FRAM repeat: matches -2..162 of consensus"
 repeat_region 33151..33174
 repeat_region /note="12 copies 2 mer tt 95% conserved"
 repeat_region 33176..33500
 repeat_region /note="LIP17 repeat: matches 1..326 of consensus"
 repeat_region 33501..33791
 repeat_region /note="HERV17 repeat: matches 8244..8523 of consensus"
 repeat_region 33811..38137
 repeat_region /note="HERV17 repeat: matches 1651..5933 of consensus"
 repeat_region 38135..39957
 repeat_region /note="HERV17 repeat: matches 1..1836 of consensus"
 repeat_region 39958..40489
 repeat_region /note="LIP17 repeat: matches 257..780 of consensus"
 repeat_region 40490..41321
 repeat_region /note="LIME3A repeat: matches 5261..6131 of consensus"
 repeat_region 41354..41655
 repeat_region /note="AluY repeat: matches 1..302 of consensus"
 repeat_region 41971..42279
 repeat_region /note="AluX repeat: matches 1..303 of consensus"
 repeat_region 43255..43365
 repeat_region /note="LIM4 repeat: matches 2169..2266 of consensus"
 repeat_region 43366..43774
 repeat_region /note="MSTA repeat: matches 1..426 of consensus"
 repeat_region 43775..44133
 repeat_region /note="LIM4 repeat: matches 2266..2691 of consensus"
 repeat_region 44140..44339
 repeat_region /note="LIM4 repeat: matches 6093..6289 of consensus"
 repeat_region 46709..47216
 repeat_region /note="LIM8 repeat: matches 5658..6169 of consensus"
 repeat_region 47217..47431
 repeat_region /note="LIM4 repeat: matches 4637..4856 of consensus"
 repeat_region 47487..47839
 repeat_region /note="MLT2E repeat: matches 8..345 of consensus"
 repeat_region 47848..48031
 repeat_region /note="LIM4 repeat: matches 4451..4645 of consensus"
 repeat_region 48017..49417
 repeat_region /note="LIMEC repeat: matches 2373..3522 of consensus"
 repeat_region 49436..50712
 repeat_region /note="LIM2D repeat: matches 5056..6333 of consensus"
 misc_feature complement(50811..51360)
 repeat_region /note="match: GSS: Em:AQ536659"
 repeat_region 52289..56182
 repeat_region /note="LIP3 repeat: matches 2255..6146 of consensus"
 repeat_region 56710..57178
 repeat_region /note="MLT1C repeat: matches 1..463 of consensus"
 repeat_region 57746..57868
 repeat_region /note="MER33 repeat: matches 202..323 of consensus"
 repeat_region 57869..58554
 repeat_region /note="MER44C repeat: matches 1..727 of consensus"
 repeat_region 58555..58752
 repeat_region /note="MER33 repeat: matches 1..202 of consensus"
 repeat_region 58890..59200
 repeat_region /note="AluSq repeat: matches 1..307 of consensus"
 repeat_region 59630..60070

misc_feature /note="LIMB3 repeat: matches 5737..6184 of consensus"
 59956..60646
 repeat_region /note="match: GSS: Em:B82923"
 61550..61582
 repeat_region /note="LIM2E repeat: matches 6133..6164 of consensus"
 61583..62103
 Query Match 86.8%; Score 551.4; DB 9; Length 128468;
 Best Local Similarity 93.9%; Pred. No. 8.8e-163;
 Matches 596; Conservative 0; Mismatches 36; Indels 3; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGATCAAACTCTAAACTA 60
 Db CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGATCAAACTCTAAACTA 33732
 QY 61 CAAATGTTCTTCAATGAGCACCAGATGGAGTCCATGACTAAAGATCCACCGTGACCC 120
 Db CAAATGTTCTTCAATGAGCACCAGATGGAGTCCATGACTAAAGATCCACCGTGACCC 33672
 QY 121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCGAG 180
 Db CT-GATCAACCTCTTAGCCCATGCTCCATGTTAATGACATTTGAAGGCAACCCCTCCGAG 33613
 QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTTAGCGGGAAGCAGTTAGACCGT 240
 Db GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTTAGCGGGAAGCAGTTAGACCGT 33553
 QY 241 CATCAGCAACCTCCCAACAGCAGCTTGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
 Db CATCAGCAACCTCCCAACAGCAGCTTGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 33493
 QY 301 AGGACTAGCTGGATTCTTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAAGGTGACT 360
 Db AGGACTAGCTGGATTCTTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAAGGTGACT 33433
 QY 361 GCATCCACCTTAAACATGGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 420
 Db GCATCTACCTTTAAACATGGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 33373
 QY 421 ACTAAATGCTAATTAGGCANAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCTG 480
 Db ACTAAGAGCTAATTAGGCANAAATAGGAGGTAAGAAATAGCCAAATCATCTATTGCTG 33313
 QY 481 AGACACACGGGAGGCAAGGATCGGATATAACCCAGGCAATTCGAGCCGGCAACGG 540
 Db AGACACACGGGAGGCAAGGATCGGATATAACCCAGGCAATTCGAGCAAGCAATGG 33253
 QY 541 CAACCCCTTTGGTCCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTCTACTCT 600
 Db CAATCCCTTTGGTCCCTCCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTCTACTCT 33195
 QY 601 ATTAATCTTGAACCTGNAAAAAAAAAAAAAAAAAAAAA 635
 Db ATTAATCTTGAACCTGNAAAAAAAAAAAAAAAAAAAAA 33160

RESULT 22
 AL139038
 LOCUS Human DNA sequence from clone Rp11-456B18 on chromosome 13,
 DEFINITION complete sequence.
 ACCESSION AL139038
 VERSION AL139038.18 GI:14800148
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 140756)
 DUNN,M.
 Direct Submission
 Submitted (14-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone

requests: clonerequest@sanger.ac.uk

On Jul 17, 2001 this sequence version replaced gi:13751261.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>

RP11-456B18 is from the library RPCT-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-456B18 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true right end of clone RP11-456B18 is at 140756 in this sequence. The true left end of clone RP11-781L6 is at 69441 in this sequence. The true right end of clone RP11-108H9 is at 100 in this sequence.

FEATURES	source
repeat_region	Location/Qualifiers 1..140756 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="13" /clone="RP11-456B18" /clone_lib="RPCT-11.2" 1830..2326 /note="LORIA repeat: matches 1..497 of consensus" 3932..3965 /note="17 copies 2 mer gt 94% conserved"
repeat_region	4632..5067 /note="JOR1b repeat: matches 1..461 of consensus"
repeat_region	5733..5893 /note="L1ME3 repeat: matches 6002..6146 of consensus"
repeat_region	6000..6083 /note="MER34 repeat: matches 454..539 of consensus"
repeat_region	6108..6171 /note="HERV23 repeat: matches 393..1056 of consensus"
repeat_region	7143..7840 /note="LTR8 repeat: matches 1..691 of consensus"
repeat_region	8059..8292 /note="LTR16A repeat: matches 203..442 of consensus"
repeat_region	8491..8686 /note="LTR29 repeat: matches 6..203 of consensus"
repeat_region	8725..8926 /note="MER4D repeat: matches 483..677 of consensus"
repeat_region	8992..9049 /note="MER4D repeat: matches 410..465 of consensus"
repeat_region	9062..9281 /note="110 copies 2 mer tt 59% conserved"
repeat_region	9398..9530 /note="MER41C repeat: matches 1..131 of consensus"
repeat_region	9623..9732 /note="LTR1 repeat: matches 675..785 of consensus"
repeat_region	9970..10333 /note="MT2D repeat: matches 1..486 of consensus"

```
repeat_region 32449..32821
/note="THE1C repeat: matches 1..371 of consensus"
repeat_region 32824..34484
/note="THE1C-internal repeat: matches 3..1651 of
consensus"
repeat_region 34465..34825
/note="THE1C repeat: matches 1..360 of consensus"
repeat_region 34804..35004
/note="MER67C repeat: matches 202..403 of consensus"
repeat_region 37922..39868
/note="LIPAL12 repeat: matches 6076..6152 of consensus"
repeat_region 40953..41258
/note="AluY repeat: matches 1..308 of consensus"
repeat_region 41279..41308
/note="15 copies 2 mer aa 86% conserved"
repeat_region 42744..42933
/note="LIMAB repeat: matches 6064..6256 of consensus"
repeat_region 42973..43275
/note="AluSX repeat: matches 1..303 of consensus"
misc_feature complement(43560..43991)
/note="match: STS: Em.HSPH07E5"
repeat_region 45196..45584
/note="MLT2B repeat: matches 2..399 of consensus"
repeat_region 45594..45767
/note="87 copies 2 mer ta 77% conserved"
repeat_region 45800..45851
/note="MLT2B repeat: matches 394..448 of consensus"
repeat_region 48118..48487
/note="THE1C repeat: matches 1..370 of consensus"
repeat_region 48708..48806
/note="AluJo/FRAM repeat: matches 197..297 of consensus"
repeat_region 49853..50271
/note="LIMEC repeat: matches 2274..2369 of consensus"
repeat_region 50637..50927
/note="LIME repeat: matches 974..1253 of consensus"
repeat_region 52250..52349
/note="MIR repeat: matches 131..228 of consensus"
repeat_region 52696..52996
/note="AluSg repeat: matches 1..293 of consensus"

Query Match 86.8%; Score 551.4; DB 9; Length 140756;
Best Local Similarity 94.3%; Pred. No. 8.9e-163;
Matches 584; Conservative 0; Mismatches 31; Indels 4; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTTTAGTTTGTCTTCAGATCAAACTGTAACACTA 60
DB 69181 CCTGTATCTTTAACTCTCTTTAGTTTGTCTTCAGATCAAACTGTAACACTA 69240

QY 61 CAAATGCTCTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 69241 CAAATGCTCTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 69300

QY 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCGAG 180
DB 69301 CTGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCGAG 69360

QY 181 GAATCTCACTGACAAACCCCTACTATGCCCAATTCAGGGGAGCAGTGTAGCGGT 240
DB 69361 GAATCTCACTGACAAACCCCTACTATGCCCAATTCAGGGGAGCAGTGTAGCGGT 69416

QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 69417 CATGCGCAACTTCCCAACAGCACTTGGGTTTTCCTGTTGAGAGGGGAGCTGAGAGAC 69476

QY 301 AGGACTAGCTGGATTTCTAGGCCCAACGAAGAATCCCTAAGCCCTAGCTGGGAAGTGACT 360
DB 69477 AGGACTAGCTGGATTTCTAGGCCCAACGAAGAATCCCTAAGCCCTAGCTGGGAAGTGACC 69536

QY 361 GCATCCACCTTAAACATGGGGCTTGCACTTACTACACCCCAACCAATCAGAGCTC 420
DB 69537 GCATCCACCTTAAACATGGGGCTTGCACTTACTACACCCCAACCAATCAGAGCTC 69596

QY 421 ACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCTG 480
```

```
DB 69597 ACTAAATGCTAACTAGGCAAAACAGGAGGTAAAGAAATAGCCAATCATCTATTGCTG 69656
QY 481 AGAGCACACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCGAGCGGCAACGG 540
DB 69657 AAAGCACACGCGGAGGACAGGATCGGATATAAACCCAGGCACTTCGAGCGGCAACAG 69716
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 69717 CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 69776
QY 601 ATTAATCTTGAACGTGA 619
DB 69777 GTTAAATCTTGAACGTGA 69795

RESULT 23
AC021774/LOCUS 167366 bp. DNA linear HTG 30-MAR-2000
DEFINITION Homo sapiens clone RP11-13C18, WORKING DRAFT SEQUENCE, 10 unordered
pieces.
AC021774 AC021774.3 GI:7341907
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 167366)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-13C18
Unpublished
2 (bases 1 to 167366)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bédaride,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Castle,A.,
Chospel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArelano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Feirreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lander,T., Lehotzky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo.A., Wu,X., Wyman,D., Ye.W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6940059.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3315
Center clone name: 13 C.18
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161819 bases at least Q40
Consensus quality: 164760 bases at least Q30
Consensus quality: 165688 bases at least Q20
Insert size: 170000; agarose-fp
```

Fri Feb 25 16:26:29 2005

Insert size: 166466; sum-of-contigs	
Quality coverage: 4.8 in Q20 bases; agarose-fp	
Quality coverage: 4.9 in Q20 bases; sum-of-contigs	

* NOTE: This is a 'working draft' sequence. It currently	
* consists of 10 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	

1	
1272: contig of 1272 bp in length	
1372: gap of 100 bp	
1373: 4088: contig of 2716 bp in length	
4089: 4188: gap of 100 bp	
4189: 6500: contig of 2312 bp in length	
6501: 6600: gap of 100 bp	
6601: 15485: contig of 885 bp in length	
15486: gap of 100 bp	
15488: 27637: contig of 12052 bp in length	
27638: 27737: gap of 100 bp	
27738: 39958: contig of 12221 bp in length	
39959: 40058: gap of 100 bp	
40059: 54618: contig of 14560 bp in length	
54619: 54718: gap of 100 bp	
54719: 82725: contig of 28007 bp in length	
82726: 82825: gap of 100 bp	
82826: 123323: contig of 40407 bp in length	
123323: 123332: gap of 100 bp	
123333: 167366: contig of 44034 bp in length.	

Location/Qualifiers	
1. .167366	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
/clone="RP11-13C18"	
/clone_lib="RPC1-11 Human Male BAC"	

1. .1272	
/note="assembly_fragment"	
1373. .4088	
/note="assembly_fragment"	
4189. .6500	
/note="assembly_fragment"	
clone_end:T7	
vector_side:right	
6601. .15485	
/note="assembly_fragment"	
15886. .27637	
/note="assembly_fragment"	
27738. .39958	
/note="assembly_fragment"	
clone_end:SP6	
vector_side:right	
40059. .54618	
/note="assembly_fragment"	
54719. .82725	
/note="assembly_fragment"	
82826. .123232	
/note="assembly_fragment"	
123333. .167366	
/note="assembly_fragment"	

ORIGIN	
Query Match 86.8%; Score 551.4; DB 2; Length 167366;	
Best Local Similarity 93.9%; Pred. No. 9.2e-163;	
Matches 596; Conservative 0; Mismatches 36; Indels 3; Gaps 2;	

1 CCCTGTATCTTTAACTCTCTTTAGTTTGTCTTCCAGATCAAACTGTAATACTA 60	
7535 CCCTGTATCTTTAACTCTCTTTAGTTTGTCTTCCAGATGAACTGTAAACTA 7476	
61 CAAATTCTTCTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120	

Db	7475	CAAAATCATTCTTCAAAATGGAGCCCGAGATGCAGTCCATGACTAAGATCCATCGACACCC	7416
Qy	121	CTGGACGGCTCTAGCCCATGCTCCGATGTTAATGACATTCAAGGCACCCCTCCCGAG	180
Db	7415	CT-GATCAACCTCTCTAGCCCATGCTCCAATGTTAATGACATTGAAGGCACCCCTCCCGAG	7357
Qy	181	GAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGAGCGGT	240
Db	7356	GAATCTCAACTGCACAAACCCCTACTATGACCCCAATTCAGCAGGAAGCAGTTAGAGCAGT	7297
Qy	241	CATCAGCAACCTCCCCAACAGACACTTGGGTTTTCTCTTTGAGAGGGGGGACTGAGAGAC	300
Db	7296	CATCAGCAACCTCCCCAACAGACACTTGGGTTTTCTCTTTGAGAGGGGGGACTGAGAGAC	7237
Qy	301	AGGACTAGCTGGATTTCTTAGGCCAACAGAAAGATCCCTAAGCCCTAGCTGGGAAAGGTGACT	360
Db	7236	AGGACTAGCTGGATTTCTTAGGCCGACTAAGAAATCCCTAAGCCCTAGCTGGGAAAGGTGACC	7177
Qy	361	GCATCCACCTCTAAACATATGGGGCTTGCACTTAGCTCACAACCCGACCAATCAGAGAGCTC	420
Db	7176	GCATCTACCTTTAAACATATGGGGCTTGCCACTTAGCTCACAACCCCAACCAATCAGAGAGCTC	7117
Qy	421	ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTGT	480
Db	7116	ACTAAGAGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCAATTGCTGT	7057
Qy	481	AGGCACAGCGGGAGGACAAAGGATCGGATATATAAACCAGGCAATTCAGCGGCAACGG	540
Db	7056	AGGCACAGTGGGAGGACAAAGATCCGGATATATAAACCAGGCAATTCAGCAGCAATGG	6997
Qy	541	CAACCCCTTTGGGTCCTCCCTTGTATGGGGCTCTGTTTTCACCTCTATTTCACCTCT	600
Db	6996	CAATCCCTTTGGGTCCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	6939
Qy	601	ATTAATCTTGCACCTGAAATGAAAAAAGAAAAA 635	
Db	6938	ATTAATCTTGCACCTGCAAAAAAAGAAAAA 6904	
RESULT 24			
AL607153	183499 bp	DNA	linear
LOCUS	Homo sapiens chromosome 13 clone RP13-276D12.		HTG 21-OCT-2001
DEFINITION			
ACCESSION	AL607153		
VERSION	AL607153.3 GI:15982188		
KEYWORDS	HTG; HTGS PHASE2; HTGS_CANCELLED.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 Lovell, J.		
AUTHORS	Direct Submission		
TITLE	Submitted (18-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,		
JOURNAL	CB10 USA, UK. E-mail enquiries: humquery@sanger.ac.uk		
requests: clonerequest@sanger.ac.uk			
On Oct 5, 2001 this sequence version replaced gi:15962724.			

Center: Sanger Centre			
Center code: SC			
Web site: http://www.sanger.ac.uk			
Contact: humquery@sanger.ac.uk			

Center project name: b276D12			

Assembly program: XGAP4; version 4.5			
Sequencing vector: plasmid; L08752; 100% of reads			
Chemistry: Dye-terminator Big Dye; 100% of reads			
Consensus quality: 183433 bases at least Q40			
Consensus quality: 183439 bases at least Q30			
Consensus quality: 183441 bases at least Q20			
Insert size: 183499; sum-of-contigs			

Insert size: 179011; 5.6% error; agarose-fp
Quality coverage: 12.84x in Q20 bases; sum-of-contigs Quality
coverage: 13.23x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1. 183499: contig of 183499 bp in length.

FEATURES

source
1. .183499
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP13-276D12"
/clone_lib="RPC1-13.2"
misc_feature
1. .183499
/note="assembly_fragment:05362"

ORIGIN

Query Match 86.8%; Score 551.4; DB 2; Length 183499;
Best Local Similarity 94.3%; Pred. No. 9.4e-163;
Matches 584; Conservative 0; Mismatches 31; Indels 4; Gaps 1;
1 CCTGTATCTTTAACTCTCTTTAGTTGTCTCTTCCAGATCAAACTGTAATACTA 60
Db 12044 CCTGTATCTTTAACTCTCTTTAGTTGTCTCTTCCAGATCAAACTGTAATACTA 12103
Qy 61 CAAATTGTTCTTCAATGGAGCCAGATGAGTGCATGACTAAGATCCACCGTGACCC 120
Db 12104 CAAATTGTTCTTCAATGGAGCCAGATGAGTGCATGACTAAGATCCACCGTGACCC 12163
Qy 121 CTGACCGGCTGTAGCCATCTCCGATGTTAATGACATTCAGGAGCCCTCCGAG 180
Db 12164 CTGACCGGCTGTAGCCATCTCCGATGTTAATGACATTCAGGAGCCCTCCGAG 12223
Qy 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTATAGACGGT 240
Db 12224 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTATAGACGGT 12279
Qy 241 CATCAGCAACCTCCCAACAGCAGTCTGGTTTCTGTTGAGAGGGGACTGAGAGAC 300
Db 12280 CATGGCCAACTCCCAACAGCAGTCTGGTTTCTGTTGAGAGGGGACTGAGAGAC 12339
Qy 301 AGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTAGTGGAGAGGTGACT 360
Db 12340 AGACTAGTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTAGTGGAGAGGTGACC 12399
Qy 361 GCATCCACCTCTAAACATGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 12400 GCATCCACCTCTAAACATGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC 12459
Qy 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGGCAATCATCTATTGCGCTG 480
Db 12460 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGGCAATCATCTATTGCGCTG 12519
Qy 481 AGACACAGCGGGAGGACAGGATCGGGATATAAACCAGGATTCGAGCCGCGACAGG 540
Db 12520 AAAGCAGCGGGAGGACAGGATCGGGATATAAACCAGGATTCGAGCCGCGACAGG 12579
Qy 541 CAACCCCTTTGGTTCCTCCCTTTGTATGGGGGCTCTGTTTTCATCTTATTTCACTCT 600
Db 12580 CAACCCCTTTGGTTCCTCCCTTTGTATGGGGGCTCTGTTTTCATCTTATTTCACTCT 12639
Qy 601 ATTAATCTTGCACACTGAA 619
Db 12640 GTTAAATCTTGCAACTGCA 12658

RESULT 25
AP002790/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AP002790 164211 bp DNA linear HTG 13-JUL-2000
Homo sapiens chromosome 18 clone RP11-732P12 map 18q21, WORKING
DRAFT SEQUENCE, 20 unordered pieces.
AP002790
HTG: HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164211)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 164,211 genomic DNA of 18q21
Published Only in DataBase (2000)
2 (bases 1 to 164211)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (11-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: Rpl1-732P12
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 154344 bases at least Q40
Consensus quality: 159150 bases at least Q30
Insert size: 162311; sum-of-contigs
Quality coverage: 5.05x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
20 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1
28606 contig of 28606 bp in length
55373 contig of 27267 bp in length
73701 contig of 17628 bp in length
92513 contig of 18712 bp in length
102171 contig of 9558 bp in length
111253 contig of 8982 bp in length
120167 contig of 8814 bp in length
127637 contig of 7370 bp in length
134267 contig of 6530 bp in length
137466 contig of 4379 bp in length
143265 contig of 4419 bp in length
147443 contig of 4278 bp in length
150342 contig of 2599 bp in length
153002 contig of 2560 bp in length
155224 contig of 2122 bp in length
156567 contig of 1243 bp in length
158332 contig of 1665 bp in length
160143 contig of 1711 bp in length
162721 contig of 2478 bp in length
164211 contig of 1390 bp in length

127738. .134267

misc_f

RESULT 26
AC064801
LOCUS
DEFINITION
AC064801
VERSION
AC064801.10 GI:22004284
KEYWORDS
HTG
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 167843)
Birtten, B., Nuebaum, C. and Lander, E.
Homo sapiens chromosome 18, clone RP11-120K19
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 167843)
Birtten, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bida, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, M., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, P., Iliiev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lechoczky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McSwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
3 (bases 1 to 167843)
Birtten, B., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tessaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
4 (bases 1 to 167843)
Birtten, B., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliiev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tessaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT
On Jul 30, 2002 this sequence version replaced gi:21700673.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8973
Center clone name: 120_K_19

Location/Qualifiers
1. .167843
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="18"
/map="18"
/clone_lib="RP11-120K19"
/clone_lib="RP11-120K19 Human Male BAC"
complement(1059..1111)
/rpt_family="MIR"
complement(1112..1424)
/rpt_family="AluX"
complement(1425..1557)
/rpt_family="MIR"
1646..1651
/note="<30 qual SINGL region"
1689..1693
/note="<30 qual SINGL region"
complement(1989..2161)
/rpt_family="MER97C"
2168..2463
/rpt_family="AluX"
complement(2489..2622)
/rpt_family="FLAM_C"
complement(3544..3967)
/rpt_family="MER97a"
4394..4633
/rpt_family="AluSx"
complement(5110..5257)
/rpt_family="MIR3"
complement(5559..5697)
/rpt_family="MER117"
6353..6769
/rpt_family="MLT1C"
6770..6846
/rpt_family="LTR16C"
complement(6847..6893)
/rpt_family="MIR"
complement(7336..8363)
/rpt_family="L1PA3"
8560..8867
/rpt_family="AluSc"
8962..9273
/rpt_family="AluSx"
11717..11744
/rpt_family="(TAAAA)n"
12041..12266
/rpt_family="MER50"


```

repeat_region 12267..12511
/rpt_family="AluSc"
repeat_region 12512..13031
/rpt_family="MER50"
repeat_region complement(14240..14530)
/rpt_family="AluJo"
repeat_region 14881..15288
/rpt_family="LTR16A1"
repeat_region 15290..15645
/rpt_family="THE1C"
repeat_region complement(15654..15794)
/rpt_family="MER5A"
repeat_region 15935..16489
/rpt_family="MLT2A2"
repeat_region complement(17190..17503)
/rpt_family="AluSx"
repeat_region complement(17524..17861)
/rpt_family="LIMEd"
repeat_region complement(17998..18430)
/rpt_family="MER65A"
repeat_region 19847..19874
/rpt_family="AT rich"
repeat_region complement(19903..20222)
/rpt_family="AluJo"
repeat_region complement(21102..21388)
/rpt_family="AluJb"
repeat_region 21410..21692
/rpt_family="L2"
repeat_region 21693..21985
/rpt_family="AluSx"
repeat_region 21986..22149
/rpt_family="L2"
repeat_region 22201..22391
/rpt_family="LTR67"
repeat_region 22657..22934
/rpt_family="AluJb"
repeat_region 22940..23062
/rpt_family="LIME3"
repeat_region 24066..24354
/rpt_family="LIME4A"
repeat_region 25031..25348
/rpt_family="AluSx"
repeat_region 25583..26009
/rpt_family="MLT1C"
repeat_region 28144..28456
/rpt_family="AluY"
repeat_region complement(28729..28847)
/rpt_family="L1MC/D"
repeat_region 29217..29290
/rpt_family="L2"
repeat_region 29341..29442
/rpt_family="MLT1I"
repeat_region 29557..29579
/rpt_family="(GAAA)n"
repeat_region complement(29806..30079)
/rpt_family="AluJb"

Query Match 86.6%; Score 550.2; DB 9; Length 167843;
Best Local Similarity 91.7%; Pred. No. 2.2e-162;
Matches 582; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCTCTTTAGTTGTCTCTTCAGAGATCAAAACTGTAAACTA 60
Db 50488 CCCTGTATCTTTAACTCTCTTTAGTTGTCTCTTCAGAGATCAAAACTGTAAACTA 50547

QY 61 CAAATGTTCTTCAATGGAGCCAGATGGAGTCAATGACTCAAGTCCACCTGGACCC 120
Db 50548 CAAATGTTCTTCAATGGAGCCAGATGGAGTCAATGACTCAAGTCCACCTGGACCC 50607

QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTGAAGCGACCCCTCCCGAG 180
Db 50608 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTGAAGCGACCCCTCCCGAG 50667

```

```

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAACAGTTAGAGCGGT 240
Db 50668 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAACAGTTAGAGCGGT 50727

QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCCCTGTTGAGAGGGGGGACTCAGAGAC 300
Db 50728 CATCAGCAACCTCCCAACAGACACTTGGGTTTCCCTGTTGAGAGGGGGGACTCAGAGAC 50787

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAAGCCCTAGCTGGGAAAGGTGACT 360
Db 50788 AGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAAGCCCTAGCTGGGAAAGGTGACT 50847

QY 361 GCATCCACCTCTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 50848 GCATCCACCTCTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 50907

QY 421 ACTAAATGCTTAATTAGGCAAAATAGAGGTAAGAAATAGCAATCATCTATTGCTG 480
Db 50908 ACTAAATGCTTAATTAGGCAAAATAGAGGTAAGAAATAGCAATCATCTATTGCTG 50967

QY 481 AGACACAGCGGGAGGACAGGATCGGATATATAACCCAGGCATTCGAGCCGGCAACGG 540
Db 50968 AGACACAGCGGGAGGACAGGATCGGATATATAACCCAGGCATTCGAGCCGGCAACGG 51027

QY 541 CAACCCCTTTGGGTCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATTTCACCTCT 600
Db 51028 CAACCCCTTTGGGTCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATTTCACCTCT 51087

QY 601 ATTAAATCTTGCACACTGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 635
Db 51088 TCAACTGCAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 51122

RESULT 27
CNS01DSH 180523 bp DNA linear PRI 20-AUG-2001
LOCUS Human chromosome 14 DNA sequence BAC R-30513 of library RPCI-11
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL121784
VERSION AL121784.5 GI:15282089
KEYWORDS HTG; HTGS; ACTIVFIN;
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 180523)
AUTHORS Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,P.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissbach,J.
Ge sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 180523)
Genoscope.
Direct Submission
Submitted (01-AUG-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
On Aug 23, 2001 this sequence version replaced gi:12001714.
----- Genom Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: Seqref@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : C-2547L24 (AC=AL135818)
Downstream BAC (overlapping the SP6 end) : R-580E16 -----
Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.78x in Q20 bases; sum-of-contigs -----

```


Overall quality chart :

Range : bases
 0 :
 1 - 9 :
 10 - 19 : 18
 20 - 29 : 87
 30 - 39 : 360
 40 - 49 : 3700
 50 - 59 : 10918
 60 - 69 : 13753
 70 - 79 : 29245
 80 - 89 : 58279
 90 - 99 : 64163

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES

source

Location/Qualifiers
 1. .180523
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone_lib="RP11-11"
 /clone="R-30513"
 1561. .1825
 /notes="matching EMBL:R74544
 R74544:RH53560
 dbSTS:ST842868
 Identified using the e-PCR software (G. Schuler)"
 69085. .69312
 /notes="matching EMBL:G03465
 R74544:RH53501
 R74544:RH3454
 dbSTS:ST81070
 Identified using the e-PCR software (G. Schuler)"
 71190. .71397
 /notes="matching EMBL:AA010373
 R74544:RH75482
 dbSTS:ST852570
 Identified using the e-PCR software (G. Schuler)"

STS

source

Location/Qualifiers
 1. .180523
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone_lib="RP11-11"
 /clone="R-30513"
 1561. .1825
 /notes="matching EMBL:R74544
 R74544:RH53560
 dbSTS:ST842868
 Identified using the e-PCR software (G. Schuler)"
 69085. .69312
 /notes="matching EMBL:G03465
 R74544:RH53501
 R74544:RH3454
 dbSTS:ST81070
 Identified using the e-PCR software (G. Schuler)"
 71190. .71397
 /notes="matching EMBL:AA010373
 R74544:RH75482
 dbSTS:ST852570
 Identified using the e-PCR software (G. Schuler)"

STS

source

Location/Qualifiers
 1. .180523
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone_lib="RP11-11"
 /clone="R-30513"
 1561. .1825
 /notes="matching EMBL:R74544
 R74544:RH53560
 dbSTS:ST842868
 Identified using the e-PCR software (G. Schuler)"
 69085. .69312
 /notes="matching EMBL:G03465
 R74544:RH53501
 R74544:RH3454
 dbSTS:ST81070
 Identified using the e-PCR software (G. Schuler)"
 71190. .71397
 /notes="matching EMBL:AA010373
 R74544:RH75482
 dbSTS:ST852570
 Identified using the e-PCR software (G. Schuler)"

ORIGIN

Query Match 86.5%; Score 549; DB 9; Length 180523;
 Best Local Similarity 92.9%; Pred. No. 5.4e-162;
 Matches 590; Conservative 0; Mismatches 35; Indels 10; Gaps 1;
 QY 1 CCCTGTATCTTAACTCTTGTAGTTTGTCTTCCAGATCAAACTGTAAACTA 60
 DB 537 CCCTTTATCTTAACTCTTGTAGTTTGTCTTCCAGATCAAACTGTAAACTA 596
 QY 61 CAAATGTTCTTCAATGGAGCCACAGATGGATGCTTCAATGATCCACCGTGGACCC 120
 DB 597 CAAATGTTCTTCAATGGAGCCACAGATGGATGCTTCAATGATCCACCGTGGACCC 656
 QY 121 CTGACCGGCTGTAGCCATGCTCCGATGTTTAAATGACATGAAGCACCCTCCCGAG 180
 DB 657 CTGACCGGCTGTAGCCATGCTCCGATGTTTAAATGACATGAAGCACCCTCCCGAG 716
 QY 181 GAAATCTCAATGACAACTCTACTATGCCCAATTCAGCGGAGCAGTGTAGAGCGT 240
 DB 717 GAAATCTCAATGACAACTCTACTATGCCCAATTCAGCGGAGCAGTGTAGAGCGT 776
 QY 241 CATCAGCACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGACTGAGAGAC 300
 DB 777 TGTGAGCACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGACTGAGAGAC 836
 QY 301 AGGACTAGCTGGATTTCTTAGGCCAACAGAGATCCCTAAGCCCTAGCTGGAGAGGTGACT 360
 DB 837 AGGACTAGCTGGATTTCTTAGGCCAACAGAGATCCCTAAGCCCTAGCTGGAGAGGTGACT 896
 QY 361 GCATCCACCTCTTAAACATGGGGCTTGAACCTAGCTCAGCCGACCAATCAGAGAGCTC 420
 DB 897 ACATCCACCTTTAAACATGGGGCTTGAACCTAGCTCAGCCGACCAATCAGAGAGCTC 956

unsure

QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAGAAATAGCAATCATCTATTGCCTG 480
 DB 957 ACTAAATGCTAATAGGCAAAATAGGAGTAAGAAATAGCAATCATCTATTGCCTG 1016
 QY 481 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCAGAGCGGCAACGG 540
 DB 1017 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCAGAGCGGCAACGG 1076
 QY 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCATCT 600
 DB 1077 CTACCTCTTTGGGTCCTCCCTTTGTATGGAGCTCAGT-----TTTCACTCT 1126
 QY 601 ATTAATCTTGAACCTGAAAAAATAAAAAA 635
 DB 1127 ATTAATCTTGAACCTGAAAAAATAAAAAA 1161

RESULT 28

AC007374 190565 bp DNA linear PRI 27-MAR-2001
 LOCUS Homo sapiens chromosome 14 clone RP11-325L17 map 14q31, complete
 DEFINITION sequence.
 AC007374
 VERSION AC007374.6 GI:13450000
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 190565)
 Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,
 Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,
 Pate, D. and Hood, L.
 Sequencing of human chromosome 15 D15S146-D15S117 region
 Unpublished
 2 (bases 1 to 190565)
 Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
 James, R., Loretz, C., Lasky, S., Madan, A., Prescott, S., Ratcliffe, A.,
 Shaffer, T. and Hood, L.
 Direct Submission
 Submitted (25-APR-1999) Multimegabase Sequencing Center, University
 of Washington, PO BOX 357730, Seattle, WA 98195, USA
 3 (bases 1 to 190565)
 Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S.,
 Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R.,
 Pate, D. and Hood, L.
 Direct Submission
 Submitted (27-MAR-2001) Multimegabase Sequencing Center, Institute
 for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
 98105, USA
 On Mar 27, 2001 this sequence version replaced gi:8247780.
 ----- Genome Center
 Center: Multimegabase Sequencing Center
 Center code: UMSC
 Web site: http://chroma.mbt.washington.edu/msg_www
 Contact: leerowen@systemsbiology.org
 ----- Summary Statistics
 Sequencing vector: pUC18; L08752
 Chemistry: Dye-terminator Big Dye; 90% of reads
 Chemistry: Dye-primer Big Dye; 10% of reads
 Assembly program: Phrap; version 0.990399

FEATURES

Location/Qualifiers
 1. .190565
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="14"
 /map="14q31"
 /clone="RP11-325L17"
 /clone_lib="RP11-325L17"
 985. .1035

```
/note="low quality data"
3865. .3870
/note="low quality data"
14875. .14880
/note="low quality data"
17060. .17095
/note="low quality data"
17488. .17535
/note="low quality data"
19433
/note="low quality data"
46474. .46580
/note="low quality data"
46695. .46700
/note="low quality data"
55163. .55242
/note="low quality data"
110860. .110875
/note="low quality data"
113613. .113617
/note="low quality data"
116230. .116782
/note="Single clone coverage. PCR and transposons tried."
117677. .117680
/note="low quality data"
117730. .117748
/note="low quality data"
118505. .118736
/note="low quality data"
144847. .144850
/note="low quality data"
152067. .152071
/note="low quality data"
158953. .158990
/note="low quality data"
177622. .177705
/note="low quality data"

ORIGIN
Query Match      86.5%; Score 549; DB 9; Length 190565;
Best Local Similarity 92.9%; Pred. No. 5.4e-162;
Matches 590; Conservative 0; Mismatches 35; Indels 10; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTGTTAGTTTGTCTCTTCAGATCAAAACTGTAACATA 60
DB 23017 CCTTTATCTTTAACTCTCTGTTAGTTTGTCTCTTCAGATCAAACTGTAACATA 23076

QY 61 CAAATTTGTTCTTCAAATGAGCACCAGATGGAGTCCATGACTAAAGATCCACCGTGGACCC 120
DB 23077 CAAATAGTTCTTCAAATGAGGCCAGATGCAGTCCATGACTAAGATCTACCATGGACCC 23136

QY 121 CTGGACCGCGCTGTCTAGCCATGCTCCGATGTTAATGACATGAAGGACCCCTCCCGAG 180
DB 23137 CTGGACCGCGCTGTCTAGCCATGCTCCGATGTTGATGACATGAAGGACCCCTCCCGAG 23196

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGCGGT 240
DB 23197 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGCGGT 23256

QY 241 CATCAGCCAACTCCCCCAACAGCAGCTTGGGTTTTCTGTGAGAGGGGGGACTGAGAGAC 300
DB 23257 TGTACGCCAACTCCCCCAACAGCAGCTTGGGTTTTCTGTGAGAGGGGGTACTGAGAGAC 23316

QY 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAGAAATCCCTAAGCCTAGCTGGGAAGTGACT 360
DB 23317 AGGACTAGCTGGATTTCTTAGGCCAACTAAGAAATCCCTAAGCCTAGCTGGGAAGTGACT 23376

QY 361 GCATCCACCTCTAAACATGGGGTTGCAACTAGCTCACACCGACCAATCAGAGAGCTC 420
DB 23377 ACATCCACCTTTAAACACAGGGGTTGCAACTAGCTCACACCGACCAATCAGAGAGCTC 23436

QY 421 ACTAAATGCTAATTAGGCCAAATAAGGAGGTAAGAAATAGCCAAATCATCTATTGCGCTG 480
```

```
Db 23437 ACTAAATGCTAATTAGGCCAAATAAGGAGGTAAGAAATAGCCAAATCATCTATCGCTG 23496
QY 481 AGACACAGCGGAGGAGGACAGGATCGGATATAAACCCAGGCAATTCGAGCCGCAACGG 540
DB 23497 AGACACAGCAGGAGGAGGACAAATGATCGGATATAAACCCAGGCAATTCGAGCCAGCATGG 23556
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTTCACCTCTATTTCACTCT 600
DB 23557 CTACCTCTTTGGGTCCCTCCCTTTGTATGGGAGCTCAGT-----TTTCACTCT 23606
QY 601 ATTAAATCTTGCAACTCAAAAAAATAAAAAA 635
DB 23607 ATTAAATCTTGCAACTGCAAAAAACAAACAAACAA 23641

RESULT 29
CNS01DVH/c 203777 bp DNA linear PRI 04-MAY-2001
LOCUS Human chromosome 14 DNA sequence BAC C-2547L24 of library Caltech-D
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL135818
VERSION AL135818.3 GI:8217905
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 203777)
Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissenbach,J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 203777)
Genoscope.
Direct Submission
Submitted (04-MAY-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Jun 3, 2000 this sequence version replaced gi:7406569.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: Seqref@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-30513
Downstream BAC (overlapping the SP6 end) : R-895M11 (AC=AL133153)
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.44x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 : 2
1 - 9 : 29
10 - 19 : 188
20 - 29 : 490
30 - 39 : 1165
40 - 49 : 5179
50 - 59 : 5946
60 - 69 : 9708
70 - 79 : 23956
80 - 89 : 61022
90 - 99 : 96092
-----
Percentage of bases with a quality value >= 40 : 99 %.
```

Location/Qualifiers
1..203777
/organism="Homo sapiens"

FEATURES
source

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

Query Match 86.3%; Score 547.8; DB 9; Length 112544;

Best Local Similarity 92.4%; Pred. No. 1.2e-161;
Matches 587; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
QY 1 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTTCAGAAATCAAAAAGTGTAAACTA 60
Db 34048 CCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTTCAGAAATCAAAAAGTGTAAACTA 34107
QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGGAGTCCATGACTAAGATCCACCGTGCACCC 120
Db 34108 CAAATCGTTCTTCAAAATGAGACCCAGATGCATGCTCAATGACTAAGATCTACCGCGACCC 34167
QY 121 CTGACCGGCTCTAGCCCATCTCCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
Db 34168 CTGACCGGCTCTAGCCCATCTCCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 34227
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 34228 GAAATCTCAACTGCACACCCCTACTACTGTCGCAATTCAGCAGGAAGCAGTTAGAGCAGT 34287
QY 241 CATCAGCCACCTCCCAACAGCAGCTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 34288 CATCGGCCAACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 34347
QY 301 AGGACTAGCTGGATTTCTAGGCCAACGAAAGATCCCTAAGCCCTAGCTGGGAAGGTGACT 360
Db 34348 AGGACTAGCTACATTTCTTAGCCGAGTAAAGATCCCTAAGCCCTAGCTGGGAAGGTGACC 34407
QY 361 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCGACCAATCAGAGAGCTC 420
Db 34408 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCTGATCAATCAGAGAGCTC 34467
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
Db 34468 ACTAAATGCTAATAGGC-AAAACAGGAGTAAAGAAATAGCAATCATCTATTGCTGCT 34526
QY 481 AGAGCAGCGGAGGAGCAAGATCGGATATAAACCCAGGATTCGAGCGGCAACGG 540
Db 34527 ACAGCAGTGGGAGGAGCAAGATCGGATGTAAACCCAGCATTCAAGCGGCAACAG 34586
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
Db 34587 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 34646
QY 601 ATTAATCTTGCACCTGAAAAAAGAAAAA 635
Db 34647 ATTAATCTTGCACCTGAAAAAAGAAAAA 34681
RESULT 31
AC009443
LOCUS Homo sapiens clone RP11-1L20, WORKING DRAFT SEQUENCE, 6 unordered
DEFINITION AC009443 192178 bp DNA linear HTG 10-NOV-2000
pieces.
ACCESSION AC009443
VERSION AC009443.4 GI:11136808
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 192178)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
Karatas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P.,
REFERENCE 1 (bases 1 to 192178)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone RP11-1L20
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 192178)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
Karatas, A., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P.,

TITLE
JOURNAL

COMMENT

Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Mellor, J., Mella, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Nilof, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Testaye, S., Torriella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
Direct Submission
Submitted (22-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 10, 2000 this sequence version replaced gi:7229764.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1480
Center clone name: 1.L.20
----- Summary Statistics
Sequencing vector: M13; W77815; 52% of reads
Sequencing vector: Plasmid; n/a; 48% of reads
Chemistry: Dye-primer-amersham; 2% of reads
Chemistry: Dye-terminator Big Dye; 98% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 189361 bases at least Q40
Consensus quality: 190532 bases at least Q30
Consensus quality: 191001 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 191678; sum-of-contigs
Quality coverage: 14.0 in Q20 bases.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES
source

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

Location/Qualifiers
1. 192178
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-1L20"
/clone_lib="RP11-1L Human Male BAC"
1. 7649
/note="assembly_fragment"
clone_end:SP6
vector_side:left
7750..9020
/note="assembly_fragment"
9121..10698
/note="assembly_fragment"
10799..12208
/note="assembly_fragment"
12309..46811
/note="assembly_fragment"
46912..192178

```

/Note="assembly_fragment
clone_end:17
vector_side:right"

ORIGIN
Query Match      86.2%; Score 547.6; DB 2; Length 192178;
Best Local Similarity 93.4%; Pred. No. 1.5e-161;
Matches 594; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 1 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
DB 179244 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 179303

QY 61 CAAATTTCTTCAAAATGAGACACAGATGGAGTCATGAC-TAAGATCCACCGTGGACC 119
DB 179304 CAAATCGTTTCTTCAAAATGAGACACAGATGGAGTCATGAC-TAAGATCCACCGTGGACC 179363

QY 120 CTTGACCGGCTGTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGA 179
DB 179364 CTTGACCGGCTGTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGA 179423

QY 180 GGAATATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGAGTTAGAGCGG 239
DB 179424 GGAATATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGAGTTAGAGCGG 179483

QY 240 TCATCAGCAACCTCCCAACAGCACTTGGGTTTCTTCTTCCAGAGGGGAGTCTGAGAGA 299
DB 179484 TCATCAGCAACCTCCCAACAGCACTTGGGTTTCTTCTTCCAGAGGGGAGTCTGAGAGA 179543

QY 300 CAGGACTAGCTGATTTCTTAGGCCAAACGAGATCCCTTAAGCTAGCTGGGAAGGTGAC 359
DB 179544 CAGGACTAGCTGATTTCTTAGGCCAAACGAGATCCCTTAAGCTAGCTGGGAAGGTGAC 179603

QY 360 TGCATCCACCTCTAAACATGGGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCT 419
DB 179604 TGCATCCACCTCTAAACATGGGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCT 179661

QY 420 CACTAAATGCTTAATTTAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCT 479
DB 179662 CACTAAATGCTTAATTTAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATCTATTGCT 179721

QY 480 GAGACACAGCGGGGACAGGATCGGATATATAACCCAGGCAATTCGAGCGGCAAG 539
DB 179722 GAGACACAGCGGGGACAGGATCGGATATATAACCCAGGCAATTCGAGCGGCAAG 179781

QY 540 GCAACCCCTTTGGGTCCTCTTGTATGGGCTCTGTTTTCACCTCTATTTCACCTC 599
DB 179782 GCAACCCCTTTGGGTCCTCTTGTATGGGCTCTGTTTTCACCTCTATTTCACCTC 179841

QY 600 TATTAATCTTCAACTGAATGAAAAAAGAAAAA 635
DB 179842 TATTAATCTTGAACCTTCAAAAAAAGAAAAAGATA 179877

RESULT 32
CNS05TF5 192899 bp DNA linear PRI 30-APR-2001
LOCUS Human chromosome 14 DNA sequence BAC R-784H12 of library RPCI-11
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL365295
VERSION AL365295.4 GI:13928046
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 192899)
AUTHORS Heilig, R., Petit, J.L., Vico V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Brule, T., Desbarrats, V., Cruaud, C.,
Gyapay, G., Saurin, W. and Weissbach, J.
TITLE Sequencing of the human chromosome 14
JOURNAL Unpublished

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

2 (bases 1 to 192899)
Genoscope.
Direct Submission
Submitted (30-APR-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
On May 2, 2001 this sequence version replaced gi:13449984.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: sequef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 6.24x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 :
1 - 9 :
10 - 19 : 5
20 - 29 : 28
30 - 39 : 267
40 - 49 : 2595
50 - 59 : 4214
60 - 69 : 9526
70 - 79 : 27497
80 - 89 : 66134
90 - 99 : 82633
-----
Percentage of bases with a quality value >= 40 : 99 %.
-----
Location/Qualifiers
1..192899
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-784H12"
/clone_lib="RPCI-11"
-----
FEATURES
source
ORIGIN
Query Match 86.2%; Score 547.6; DB 9; Length 192899;
Best Local Similarity 93.4%; Pred. No. 1.5e-161;
Matches 594; Conservative 0; Mismatches 39; Indels 3; Gaps 2;
QY 1 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
DB 14879 CCCTGTATCTTTAACTCCTCTTAAAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 14938
QY 61 CAAATTTCTTCAAAATGAGACACAGATGGAGTCATGAC-TAAGATCCACCGTGGACC 119
DB 14939 CAAATCGTTTCTTCAAAATGAGACACAGATGGAGTCATGAC-TAAGATCCACCGTGGACC 14998
QY 120 CTTGACCGGCTGTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGA 179
DB 14999 CTTGACCGGCTGTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGA 15058
QY 180 GGAATATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGAGTTAGAGCGG 239
DB 15059 GGAATATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGAGTTAGAGCGG 15118
QY 240 TCATCAGCAACCTCCCAACAGCACTTGGGTTTCTTCTTCCAGAGGGGAGTCTGAGAGA 299
DB 15119 TCATCAGCAACCTCCCAACAGCACTTGGGTTTCTTCTTCCAGAGGGGAGTCTGAGAGA 15178
QY 300 CAGGACTAGCTGATTTCTTAGGCCAAACGAGATCCCTTAAGCTAGCTGGGAAGGTGAC 359
DB 15179 CAGGACTAGCTGATTTCTTAGGCCAAACGAGATCCCTTAAGCTAGCTGGGAAGGTGAC 15238
QY 360 TGCATCCACCTCTAAACATGGGGCTTGCACCTAGCTACACCCGACCAATCAGAGAGCT 419

```

```

Db 15239 TGCATCCACCTTTAAACACGGGCTTGCACTTAGCTCACACCAACCAATC--AGAGCT 15296
QY 420 CACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCT 479
Db 15297 CACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCT 15356
QY 480 GAGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACG 539
Db 15357 GAGAGCACAGGAGGAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACG 15416
QY 540 GCAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTTATTCACT 599
Db 15417 GCAACCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGTTTCACTTATTCACT 15476
QY 600 TATTAATCTTGCACTGAATAAAAAAAGAAAAA 635
Db 15477 TATTAATCTTGCACTTCAAAAAAAGAAAAA 15512

```

```

RESULT 33
AC068492 172281 bp DNA linear PRI 29-OCT-2002
LOCUS Homo sapiens BAC clone RP11-809C23 from 2, complete sequence.
DEFINITION AC068492
ACCESSION AC068492
VERSION AC068492.2 GI:14010945
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 172281)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 172281)
AUTHORS Lamar, B., Abbott, A. and Ureta, M.
TITLE The sequence of Homo sapiens BAC clone RP11-809C23
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 172281)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 172281)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 172281)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 6 (bases 1 to 172281)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 9, 2001 this sequence version replaced gi:7677994.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0809C23
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-809C23; actual end is at base position 172281 of RP11-809C23.

FEATURES	Location/Qualifiers
source	1..172281
	/organism="Homo sapiens"
	/mol_type="Genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-809C23"
	/clone_lib="RPCI-11"
repeat_region	1161..1218
	/rpt_family="MIR"
repeat_region	1245..1302
	/rpt_family="MIR"
repeat_region	1313..1407
	/rpt_family="MIR"
repeat_region	2612..2712
	/rpt_family="MaLR"
repeat_region	2868..3080
	/rpt_family="ERV1"
repeat_region	3088..3175
	/rpt_family="MER1_type"
repeat_region	3872..3997
	/rpt_family="MIR"
repeat_region	4265..4436
	/rpt_family="MIR"
repeat_region	5913..6082
	/rpt_family="MER1_type"
repeat_region	6124..6274
	/rpt_family="L2"
repeat_region	6855..7214
	/rpt_family="MaLR"
repeat_region	7808..7956
	/rpt_family="MIR"
repeat_region	7990..8112
	/rpt_family="L2"
repeat_region	8188..8540
	/rpt_family="MaLR"
repeat_region	8659..8760
	/rpt_family="MER53"
repeat_region	9485..9630
	/rpt_family="MaLR"

JOURNAL

Submitted (05-DEC-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA

REFERENCE
AUTHORS

3 (bases 1 to 169418)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.
and Haugen, E.D.

TITLE
JOURNAL

Direct Submission
Submitted (31-MAY-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA

COMMENT

On May 31, 2002 this sequence version replaced gi:17352427.

Center: University of Washington Genome Center	3814	3835	8434	8627	512	<800
Center Code: UWGC	4550	4514	174	<800	449	<800
Web site: http://www.genome.washington.edu	13599	13502	3650	3671	3186	3034
Contact: wgchgs@u.washington.edu	1967	2060	1165	1179	1119	1142
Drafting Center: WUGSC	4453	4514	3474	3428	1113	1142
----- Project Information -----	265	<800	785	<800	11537	11468
Center project name: chr-3	7594	7642	2069	2051	11876	11468
Center clone name: RP11-66G8 (bc0186)	3677	3631	6745	6886	2248	2332
----- Summary Statistics -----	6124	5984	2015	2051	7791	8102
Sequencing vector: unknown; 62% of reads	66	<800	990	982	11650	11468
Sequencing vector: plasmid; L08752; 38% of reads	1869	1860	10194	10069	9162	9102
Chemistry: Dye-terminator ET; 66% of reads	214	<800	7207	6886	3337	3297
Chemistry: Dye-terminator Big Dye; 34% of reads	4500	4514	6300	6210	1749	1637
Assembly program: Phrap; version 0.990319	4632	4514	66	<800	704	<800
Consensus quality: 169322 bases at least Q40	2817	2875	8157	8155	542	<800
Consensus quality: 169405 bases at least Q30	305	<800	6827	6886	3662	3500
Consensus quality: 169418 bases at least Q20	3843	3835	8122	8155	239	<800
Insert size: 169418; sum-of-contigs	504	<800	1177	1179	1653	1637
Quality coverage: 12.2x in Q20 bases; sum-of-contigs	898	898	266	<800	4537	4514
----- Overlapping Sequences -----	1517	1455	936	982	6109	6059
5': RP11-420K5 (UWGC:bc0651) AC104444, 57330-bp overlap	5718	5984	78	<800	371	<800
3': RP11-453D3 (UWGC:bc0429) AC104188, 73864-bp overlap	1466	1455	2263	2285	8219	8102
----- Sequence Quality Assessment -----	2965	2875	15109	15154	1904	1918
This entry has been annotated with sequence quality	1642	1613	1213	1179	3938	3854
estimates computed by the Phrap assembly program.	6084	5984	10117	10069	409	<800
All manually edited bases have been reduced to quality zero.	4930	4849	1857	1857	1800	1918
Quality levels above 40 are expected to have less than	12	<800	1328	1314	1886	1918
1 error in 10,000 bp.	3848	3835	4583	4553	2804	2613
Base-by-base quality values are not generally visible from the	1681	1613	6464	6479	1955	1918
Genbank flat file format but are available as part	1858	1860	2865	2916	3230	3297
of this entry's ASN.1 file.	897	898	2362	2418	195	<800
This sequence was finished as follows unless otherwise noted:	4602	4514	2046	2051	7172	7125
all regions were either double-stranded or sequenced with an	2839	2875	6729	6886	1173	1142
alternate chemistry or covered by high quality data (i.e., Phred	649	<800	2952	2916	1572	1637
quality >= 30); an attempt was made to resolve all sequencing	2131	2060	1712	1683	514	<800
problems, such as compressions and repeats; all regions were	2302	2318	309	<800	2058	2116
covered by at least one plasmid subclone or more than one M13						
subclone; and the assembly was confirmed by restriction digest.						
----- Sequence Validation -----						
This sequence has been validated by Multiple Complete Digest						
fingerprinting. Comparison of the experimentally derived digest						
fragments with sequence-predicted fragments is given below.						
The electronically-digested sequence consists of both insert and						
vector, in order to accurately represent the entire circular BAC.						
Small fragments below a variable cutoff (approximately 400-800 bp)						
are not resolved in the fingerprint and hence do not appear						
in the table. There are no significant remaining discrepancies						
between the experimental and predicted values. Uniquely ordered						
fragments are separated by dashed lines.						

ECORI						
BglII						
HindIII						
SeqDerMap FngPrnt SeqDerMap FngPrnt						

8696 8873 8104 8155 459 <800						

6 <800 2067 2051 6382 6462						

QY	481	AGAGCACAGCGGAGGACAAAGGATCGGGATATATAAACCCAGGCATTCGAGCGCGCAACGG	540
Db	160039	AGAGCACAGTGGGTGGGACAAAGGATCGGGATATATAAACCCAGGCATTTGAACTGGCAATGG	159980
QY	541	CAACCCCTTTGGGTCCCTCCCTTCCTTTGATGGGGCTCTGTGTTTTCACCTCTATTTTTCACCTCT	600
Db	159979	CACCCCTTTGGGTCTCTCTCCCTTTGATGGGAGCTCTGTGTTTTCACCTCTATTTTTCACCTCT	159920
QY	601	ATTAATAATCTTGCACACTGAAAAAAGAAAAAAGAAAAA 634	
Db	159919	ATTAATAATCTTGCACACTCCAAAAAAGAAAAAAGAAAAA 159886	
RESULT 35			
AC104444			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

Query Match	86.2%	Score 547.2	DB 9	Length 169418
Best Local Similarity	92.6%	Pred. No. 2e-164		
Matches 587	Conservative 0	Mismatches 43	Indels 4	Gaps 1
QY 1	CCCTGTATCTTAAACCTCCTTGGTAACTTTGTCTCTTCCAGAAATCAAAACTGTAAAACTTA	60		
DB 150515	CCCTGTATCTTAAATCTCTTATTAACTTTGTTTCTTCCAGAAATCGAAGCTGTAAAACTTA	160456		
QY 61	CAAAATGTTCTTCAAAATGAGACACAGATGAGTCCATGACTAAGATCCACCGTGAACCC	120		
DB 160455	CAAAATCGTTCTTCAAAATGAGACCCAGATGAGTCCATGACTAAGATCTACTGCAAGACCC	160396		
QY 121	CTGGACCGGCTGCTAGGCCATGCTCGAGTGTAAATGACATTGAAGGCACCCCTCCCGAG	180		
DB 160395	CTGGACTGGCTGCTAGGCCCATGCTTGTAGTGTAAATGACATCGAAGGCACCCCTCCCGAG	160336		
QY 181	GAATPTCTCAACTGCACAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGCGGT	240		
DB 160335	GAATPTCTCAACTGCACAACCACTACTACACCCCAATTCAGCAAGGAAGCAGTTAGAGTGT	160276		
QY 241	CATCAGCCAACTCCCCCAACAGCACTTTGGGTTTTCTGTTGAGAGGGGGGACCTGACAGAC	300		
DB 160275	CATCGGCCAACTCCCCCAACAGCACTTTGGGTTTTCTGTTGAGAGGGGGGATTTGAGAGAT	160216		
QY 301	AGGACTAGCTGGATTTCTTAGGCCAAAGAAATCCCTTAAGCCTTAGCTGGGAAGTGTACT	360		
DB 160215	AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTTAGCCCTTAGCTGGGAAGTGTACC	160156		
QY 361	GCATCCACTCTAAACATGGGCTTGCACCTTACCTCAGCCCGAACCAATCAGAGAGCTC	420		
DB 160155	ACATCCACCTTTAAACATGGGCTTTGCACCTTACCTCAGCCCGAACCAATCAGAGAGCTC	160100		
QY 421	ACTAAATGCTAAATTAGGCAAAAAATGAGAGTAAAGAAATAGCCAAATCATCTATTGGCTG	480		
DB 160099	ACTAAATGCTAAATTAGGCAAAAAATGAGAGTAAAGAAATAGCCAAATCATCTATTGGCTG	160040		

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.

				HindIII			
				EcoRI			
SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt
4139	4075	8696	8846	5016	4986		
2067	2050	6	<800	6382	6692		
4733	4896	4633	4582	512	<800		
8157	8220	4499	4384	449	<800		
66	<800	214	<800	8852	8947		
6300	6257	1869	1869	11649	11652		
7206	7163	66	<800	7791	7977		
10189	10145	6123	5997	2243	2336		
990	991	3677	3678	11879	11652		
2018	2050	7589	7598	11549	11652		
6744	6810	265	<800	1113	1110		
2069	2050	4453	4384	1119	1110		
784	772	1970	1952	2300	2336		
3476	3398	13599	13498	2080	2102		
1165	1150	4562	4582	6024	5884		
3662	3684	3815	3678	2278	2336		
174	<800	10	<800	1839	1862		
5486	5446	599	<800	4646	4692		
1285	1285	5326	5404	1421	1401		
4875	4896	2528	2566	3376	3400		
350	<800	10560	10460	2359	2336		
327	<800	1581	1540	1906	1862		
6737	6810	2287	2274	863	899		
173	<800	516	<800	3833	3828		

FEATURES

source
Location/Qualifiers
1. .161054
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-420K5"
/clone_lib="RPCI human BAC library 11"

Query Match 86.1%; Score 546.6; DB 9; Length 161054;
Best Local Similarity 92.4%; Pred. No. 3.1e-161;
Matches 587; Conservative 0; Mismatches 44; Indels 4; Gaps 1;

Fri Feb 25 16:26:29 2005

QY 1 CCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
 Db 112631 CCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGAAATCGAAGCTGTAAACTA 112690
 QY 61 CAAATGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 112691 CAAATCAATCTTCAAAATGGAGCCAGATGAGTCCATGACTAAGATCTACTGAGACGCC 112750
 QY 121 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 180
 Db 112751 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 112810
 QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGGAAGCAGTTAGAGCGGT 240
 Db 112811 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGGAAGCAGTTAGAGTGGT 112870
 QY 241 CATCAGCCAACTCCCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
 Db 112871 CATCAGCCAACTCCCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 112930
 QY 301 AGGACTAGCTGATTTCTTAGCCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 112931 AGGACTAGCTGATTTCTTAGCCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 112990
 QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTTAGCTCAGCCGACCAATCAGAGAGCTC 420
 Db 112991 ACATCCACCTTTAAACATGGGCTTGCAACTTTAGCTCAGCCGACCAATCAGAGAGCTC 113046
 QY 421 ACTAAATGCTAATTAGGCAGAAATAGGAGGTAAGAAATAGCCCAATCATCTATTGCGCTG 480
 Db 113047 ACTAAATGCTAATTAGGCAGAAATAGGAGGTAAGAAATAGCCCAATCATCTATTGCGCTG 113106
 QY 481 AGAGCAGCGGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACGG 540
 Db 113107 AGAGCAGCGGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACGG 113166
 QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTCACTCT 600
 Db 113167 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTCACTCT 113226
 QY 601 ATTAATCTTGCACCTGAAAAAAGAAAAA 635
 Db 113227 ATTAATCTTGCACCTGAAAAAAGAAAAA 113261

RESULT 36
 AX478550
 LOCUS AX478550 2074 bp DNA linear PAT 12-AUG-2002
 DEFINITION Sequence 4 from Patent WO0248362.
 ACCESSION AX478550
 VERSION AX478550.1 GI:22217319
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 Ramkumar,J. and Arvizu,C.
 Embryogenesis associated proteins
 TITLE Patent: WO 0248362-A 4 20-JUN-2002;
 JOURNAL Incyte Genomics, Inc. (US)
 FEATURES
 source
 1. 2074
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7477736CB1"
 ORIGIN
 Query Match 86.0%; Score 546.4; DB 6; Length 2074;
 Best Local Similarity 93.8%; Pred. No. 1.7e-161;
 Matches 579; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 1 CCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
 Db 1459 CCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 1518
 QY 61 CAAATGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 1519 CAAATGTTCTTCAAAATGGAGCCAGATGAGTCCATGACTAATCTTACACGAGGCC 1578
 QY 121 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 180
 Db 1579 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATCGAAGGCACTCTCTCCCAAG 1638
 QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGGAAGCAGTTAGAGCGGT 240
 Db 1639 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGGAAGCAGTTAGAGCGGT 1698
 QY 241 CATCAGCCAACTCCCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
 Db 1699 CGTCAGTCAACCTCCCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 1758
 QY 301 AGGACTAGCTGATTTCTTAGCCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 1759 AGGACTAGCTGATTTCTTAGCCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 1818
 QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTTAGCTCAGCCGACCAATCAGAGAGCTC 420
 Db 1819 GCGTCCACCTTTAAACAGCGGGCTTGCAACTTTAGCTCAGCCGACCAATCAGAGAGCTC 1878
 QY 421 ACTAAATGCTAATTAGGCAGAAATAGGAGGTAAGAAATAGCCCAATCATCTATTGCGCTG 480
 Db 1879 ACTAAATGCTAATTAGGCAGAAATAGGAGGTAAGAAATAGCCCAATCATCTATTGCGCTG 1938
 QY 481 AGAGCAGCGGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACGG 540
 Db 1939 AGAGCAGCGGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACGG 1997
 QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTCACTCT 600
 Db 1998 CAACCGCTTTGGTCCCTCCCTTTGATGGAGCTCTGTTTCACTCTATTCACTCT 2057
 QY 601 ATTAATCTTGCACCTG 617
 Db 2058 ATTAATCTTGCACCTG 2074

RESULT 37
 CNS06C7Q/c
 LOCUS CNS06C7Q 169234 bp DNA linear PRI 28-APR-2001
 DEFINITION Human chromosome 14 DNA sequence BAC R-21A20 of library RPCI-11 from chromosome 14 of Homo sapiens (Human), complete sequence.
 ACCESSION AL390799
 VERSION AL390799.4 GI:13516219
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 169234)
 Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P., Brothier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F., Levy,M., Eckenberg,R., Bruls,T., DeBerardinis,V., Cruaud,C., Gyapay,G., Saurin,W. and Weissenbach,J.
 TITLE Sequencing of the human chromosome 14
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 169234)
 Genoscope.
 Direct Submission
 TITLE Submitted (26-APR-2001) Genoscope - Centre National de Sequencage : Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 JOURNAL On Apr 2, 2001 this sequence version replaced gi:13016608.
 COMMENT ----- Genome Center

Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: <http://www.genoscope.cns.fr/>
Contact: SeqRef@genoscope.cns.fr

The following BAC sequence is oriented from the T7 to the SP6 end.

----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 8.21x in Q20 bases; sum-of-contigs

Overall quality chart :

Range : bases

0 - 9 : 5

10 - 19 : 24

20 - 29 : 84

30 - 39 : 1610

40 - 49 : 3705

50 - 59 : 5060

60 - 69 : 13612

70 - 79 : 43906

80 - 89 : 101228

90 - 99 : 101228

Percentage of bases with a quality value >= 40 : 99 %.

Location/Qualifiers

1. .169234

/organism="Homo sapiens"

/mol type="genomic DNA"

/db xref="taxon:9606"

/chromosomes="14"

/clone="R-21A20"

/clone lib="RPC1-11"

136833. .137011

/notes="matching EMBL:Z52518"

RHdb:RH31311

dbSTS:STS28504

Identified using the e-PCR software (G. Schuler)"

160586. .160738

/notes="matching EMBL:T77897"

RHdb:RH44683

dbSTS:STS37751

Identified using the e-PCR software (G. Schuler)"

Query Match 86.0%; Score 546; DB 9; Length 169234;

Best Local Similarity 93.2%; Pred. No. 4.8e-161;

Matches 593; Conservative 0; Mismatches 40; Indels 3; Gaps 2;

Qy 1 CCCTGTATCTTTAACTCTCTGTGTTAGTTGCTCTTCCAGAAATCAAACTGTAAACTA 60

Db 12949 CCCTGTATCTTTAACTCTCTGTGTTAGTTGCTCTTCCAGAAATCAAACTGTAAACTA 12890

Qy 61 CAAATTGTTCTTCAATGGAGCACCACATGGAGTCCATGAC-TAAGATCCACCGTGACC 119

Db 12889 CAAATGTTCTTCAATGGAGCACCACATGGAGTCCATGAC-TAAGATCCACCGTGACC 12830

Qy 120 CCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACCCCTCCCGA 179

Db 12829 CCTAGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACCCCTCCCGA 12770

Qy 180 GGAATCTCACTGACCAACCCCTATATGCCCCCAATTCAGCGGGAAGAGTTAGCGG 239

Db 12769 GGAATCTCACTGACCAACCCCTATATGCCCCCAATTCAGCGGGAAGAGTTAGCGG 12710

Qy 240 TCATCAGCACTCTCCCAACAGCACTTGTTTCTGTTGAGAGGGGAGTACGAGA 299

Db 12709 TCGTCGATCACTCTCCCAACAGCACTTGTTTCTGTTGAGAGGGGAGTACGAGA 12650

Qy 300 CAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCCTAGCTGGAGGTGAC 359

Db 12649 CAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCCTAGCTGGAGGTGAC 12590

FEATURES

source

STS

STS

ORIGIN

Query Match 86.0%; Score 546; DB 9; Length 169234;
Best Local Similarity 93.2%; Pred. No. 4.8e-161;
Matches 593; Conservative 0; Mismatches 40; Indels 3; Gaps 2;
Qy 1 CCCTGTATCTTTAACTCTCTGTGTTAGTTGCTCTTCCAGAAATCAAACTGTAAACTA 60
Db 12949 CCCTGTATCTTTAACTCTCTGTGTTAGTTGCTCTTCCAGAAATCAAACTGTAAACTA 12890
Qy 61 CAAATTGTTCTTCAATGGAGCACCACATGGAGTCCATGAC-TAAGATCCACCGTGACC 119
Db 12889 CAAATGTTCTTCAATGGAGCACCACATGGAGTCCATGAC-TAAGATCCACCGTGACC 12830
Qy 120 CCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACCCCTCCCGA 179
Db 12829 CCTAGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTAAGGACCCCTCCCGA 12770
Qy 180 GGAATCTCACTGACCAACCCCTATATGCCCCCAATTCAGCGGGAAGAGTTAGCGG 239
Db 12769 GGAATCTCACTGACCAACCCCTATATGCCCCCAATTCAGCGGGAAGAGTTAGCGG 12710
Qy 240 TCATCAGCACTCTCCCAACAGCACTTGTTTCTGTTGAGAGGGGAGTACGAGA 299
Db 12709 TCGTCGATCACTCTCCCAACAGCACTTGTTTCTGTTGAGAGGGGAGTACGAGA 12650
Qy 300 CAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCCTAGCTGGAGGTGAC 359
Db 12649 CAGGACTAGCTGGATTTCTTAGGCCAACGAGAAATCCCTAAGCCCTAGCTGGAGGTGAC 12590

Qy 360 TGCATCCACCTCTAAACATGGGGCTTGAACCTTAGCTCACCCGACCAATCAGAGCT 419
Db 12589 TGCATCCACCTTTAAACACGGGGCTTGAACCTTAGCTCACCCGACCAATCAGAGCT 12532
Qy 420 CACTAAATGCTAATAGGCAAAATAGGAGTAAGAATAAGCAATCATCTATTGCTT 479
Db 12531 CACTAAATGCTAATAGGCAAAATAGGAGTAAGAATAAGCAATCATCTATTGCTT 12472
Qy 480 GAGACACAGCGGGAGGAGGACAGGATCGGATATAAACCAGGCATTCGAGCGGCAACG 539
Db 12471 GAGACACAGCGGGAGGAGGACAGGATCGGATATAAACCAGGCATTCGAGCGGCAACG 12412
Qy 540 GCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACTCTATTTCACCTC 599
Db 12411 GCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACTCTATTTCACCTC 12352
Qy 600 TATTAAATCTTGAACACTGAAAAAAGAAAAA 635
Db 12351 TATTAAATCTTGAACACTTCAAAAAAAGAAAAAGATA 12316

RESULT 38
BD136199
LOCUS
DEFINITION
BD136199 1329 bp DNA linear PAT 18-SBP-2002
Retroviral nucleic material and nucleotide fragments, in
particular, associated with multiple sclerosis and/or rheumatoid
arthritis, for diagnostic, prophylactic and therapeutic uses.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
BD136199
BD136199.1 GI:23231144
JP 2002509437-A/9.
unidentified
unidentified
unclassified.
1 (bases 1 to 1329)
Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F.,
Perron,H. and Mandrand,B.
Retroviral nucleic material and nucleotide fragments, in
particular, associated with multiple sclerosis and/or rheumatoid
arthritis, for diagnostic, prophylactic and therapeutic uses
Patent: JP 2002509437-A 9 26-MAR-2002;
BIO MERIEUX
OS Unidentified
PN JP 2002509437-A/9
PD 26-MAR-2002
PF 07-JUL-1998 JP 1999508255
PR 07-JUL-1997 FR 97/08816
PI GLAUCIA PARAHNOS BACCALA, FLORENCE KOMURIAN PRADEL, FREDERIC PI
BEDIN,
PI MIREILLE SODOYER, CATHERINE OTT, FRANCOIS MALLET, HERVE PERRON,
PI BERNARD MANDRAND
PC C12N15/48,C12Q1/70,C07K14/15,A61K31/70
CC Strandedness: Single;
CC Topology: Linear;
CC Retroviral nucleic material and nucleotide fragments, in CC
CC associated with multiple sclerosis and/or
rheumatoid arthritis,
for
diagnostic, prophylactic and therapeutic uses FH Key
CC Location/Qualifiers
FT source 1. .1329
FT Location/Qualifiers
/organism='Unidentified'.
1. .1329
/organism='unidentified'
/mol type='genomic DNA'
/db_xref='taxon:32644'

FEATURES

source

ORIGIN

Query Match 85.8%; Score 545; DB 6; Length 1329;
Best Local Similarity 92.4%; Pred. No. 4.3e-161;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60
DB 257 CCCTGTATCTTCAACTCTCTGTTAAAGTTTGTCTCTCCAGAAATGAAAGCTGTAAAGCTA 316
QY 61 CAAATGTTCTTCAAAATGGAGACACAGATGGAGTCCATGACTAAAGACTCCACCGTGGAGCC 120
DB 317 CAAATAGTTCTTCAAAATGGAAACCCAGATGAGTCCATGACTTAAATCTACCGTGGAGCC 376
QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGGACACCCCTCCCGAG 180
DB 377 CTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAG 436
QY 181 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTCAGCGGGAAGCAAGTTAGAGCGGT 240
DB 437 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTCAGTGAAGCAGTTAGAGCAGT 496
QY 241 CATCAGCAACTCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGACTGAGAGAC 300
DB 497 TGTGAGCCAACTCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGAC 556
QY 301 AGGACTAGCTGATTTCTTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAAGTGTACT 360
DB 557 AGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGGGAAGTGTACC 616
QY 361 GCATCCACTCTAAACATGGGCTTGCACACTTGTAGCTCACCCGACCAATCAGAGAGCTC 420
DB 617 GCATCCATCTTTAAACATGGGCTTGCACACTTGTAGCTCACCCGACCAATCAGAGAGCTC 676
QY 421 ACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
DB 677 ACTAAATCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCCAATCATCTATTGCGCTG 736
QY 481 AGAGCACAGCGGAGGACAAGGATCGGGATATAAACCAGGCAATTCAGAGCGGCAACGG 540
DB 737 AGAGCACAGCGGAGGACAAGGATCGGGATATAAACCAGGCAATTCAGAGCGGCAACAG 796
QY 541 CAACCCCTTTGGGTCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 797 CAACCCCTTTGGGTCCTCCCTTGTATGGAGCTCTGTTTCACTCTATTTCACCTCT 856
QY 601 ATTAATCTTGAACCTGAA 619
DB 857 ATTAATCATGCACTGCA 875

RESULT 39
AR344389
LOCUS AR344389 1329 bp mRNA linear PAT 17-AUG-2003
DEFINITION Sequence 108 from patent US 6582703.
ACCESSION AR344389
VERSION AR344389.1 GI:33740330
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1329)
Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G.,
Komurian-Pradel,F., Jollivet-Reynaud,C. and Mandrand,B.
Isolated nucleotide sequences associated with multiple sclerosis or
rheumatoid arthritis and a process of detecting
JOURNAL Patent: US 6582703-A 108 24-JUN-2003;
FEATURES
Location/Qualifiers
1..1329
/organism="unknown"
/mol_type="mRNA"
ORIGIN
Query Match 85.8%; Score 545; DB 6; Length 1329;
Best Local Similarity 92.4%; Pred. No. 4.3e-161;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60

DB 257 CCCTGTATCTTCAACTCTCTGTTAAAGTTTGTCTCTCCAGAAATGAAAGCTGTAAAGCTA 316
QY 61 CAAATGTTCTTCAAAATGGAGACACAGATGGAGTCCATGACTAAAGACTCCACCGTGGAGCC 120
DB 317 CAAATAGTTCTTCAAAATGGAAACCCAGATGAGTCCATGACTTAAATCTACCGTGGAGCC 376
QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGGACACCCCTCCCGAG 180
DB 377 CTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAG 436
QY 181 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTCAGCGGGAAGCAAGTTAGAGCGGT 240
DB 437 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTCAGTGAAGCAGTTAGAGCAGT 496
QY 241 CATCAGCAACTCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGACTGAGAGAC 300
DB 497 TGTGAGCCAACTCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGAC 556
QY 301 AGGACTAGCTGATTTCTTAGGCCAAAGAAATCCCTAAGCTAGCTGGGAAGTGTACT 360
DB 557 AGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGGGAAGTGTACC 616
QY 361 GCATCCACTCTTAAACATGGGCTTGCACACTTGTAGCTCACCCGACCAATCAGAGAGCTC 420
DB 617 GCATCCATCTTTAAACATGGGCTTGCACACTTGTAGCTCACCCGACCAATCAGAGAGCTC 676
QY 421 ACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCGCTG 480
DB 677 ACTAAATCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCCAATCATCTATTGCGCTG 736
QY 481 AGAGCACAGCGGAGGACAAGGATCGGGATATAAACCAGGCAATTCAGAGCGGCAACGG 540
DB 737 AGAGCACAGCGGAGGACAAGGATCGGGATATAAACCAGGCAATTCAGAGCGGCAACAG 796
QY 541 CAACCCCTTTGGGTCCTCCCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 797 CAACCCCTTTGGGTCCTCCCTTGTATGGAGCTCTGTTTCACTCTATTTCACCTCT 856
QY 601 ATTAATCTTGAACCTGAA 619
DB 857 ATTAATCATGCACTGCA 875

RESULT 40
AX001030
LOCUS AX001030 1329 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 12 from Patent WO9902666.
ACCESSION AX001030
VERSION AX001030.1 GI:7241266
KEYWORDS
SOURCE unidentified
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 1329)
Ott,C. and Bedin,F.
RETROVIRAL NUCLEIC MATERIAL AND NUCLEOTIDE FRAGMENTS, IN PARTICULAR
ASSOCIATED WITH MULTIPLE SCLEROSIS AND/OR RHEUMATOID ARTHRITIS, FOR
DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC USES
JOURNAL Patent: WO 9902666-A 12 21-JAN-1999;
BIO MERIEUX (FR); OTT CATHERINE (FR)
FEATURES
Location/Qualifiers
1..1329
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 85.8%; Score 545; DB 6; Length 1329;
Best Local Similarity 92.4%; Pred. No. 4.3e-161;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCCTCTGTTAAAGTTTGTCTCTCCAGAAATCAAAACTGTAAACTA 60

[illegible]

Search completed: February 21, 2005, 11:52:17
Job time : 2738.62 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:25:25 ; Search time 327.546 Seconds
(without alignments)
11476.370 Million cell updates/sec

Title: US-09-319-156B-6
Perfect score: 635
Sequence: 1 cccgtatcttttaacctctt.....tgaaaaa.....635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	635	100.0	635	2 AAV43215	AAV43215 Multiple
2	635	100.0	635	2 AAX29702	Aax29702 Clone CL6
3	635	100.0	635	12 ADG14843	Adg14843 MSRV asso
4	626	98.6	635	9 ADB84397	Adb84397 MSRV-1 as
5	617.4	97.2	2030	3 AAA63826	Aaa63826 Nucleotid
6	546.4	86.0	2074	6 AAD41225	Aad41225 Human EMB
7	545	85.8	1329	2 AAV43219	Aav43219 Multiple
8	545	85.8	1329	2 AAX29704	Aax29704 Clone 5M6
9	545	85.8	1329	12 ADG14849	Adg14849 MSRV asso
10	543.4	85.6	21646	11 ACN44454	Actn44454 Human gen
11	540.2	85.1	2046	8 ACC46747	Acc46747 Human dit
12	535	84.3	1329	9 ADB84403	Adb84403 MSRV-1 as
13	533.8	84.1	2946	2 AAX77526	Aax77526 Human sec
14	533.8	84.1	2946	3 AAZ59468	Aaz59468 Human sec
15	533.8	84.1	2946	10 ADC38776	Adc38776 Human cDN
16	532.8	83.9	326014	6 AKB83296	Akb83296 Human gen
17	532.8	83.9	326014	12 ADQ94981	Adq94981 Human kin
18	532.2	83.8	2781	5 AAF55630	Aaf55630 Nucleotid
19	525.8	82.8	1136	2 AAX25660	Aax25660 Human end
20	525.8	82.8	1136	3 AAA59210	Aaa59210 3' pol ge

21	522.6	82.3	2782	2 AAX25661	Aax25661 Human end
22	522.6	82.3	2782	3 AAA59211	Aaa59211 5' non co
23	522.6	82.3	2782	5 AAH20069	Aah20069 HERV-w en
24	517.4	81.5	2930	6 AAD24195	Aad24195 Human syn
25	516.2	81.3	5693	6 ABL61744	Ab161744 Colon ade
26	511.4	80.5	7582	2 AAX25665	Aax25665 Complete
27	511.4	80.5	7582	3 AAA59215	Aaa59215 Human end
28	510.6	80.4	46340	3 ABN97978	Abn97978 Human ret
29	510.6	80.4	161334	11 ACN44334	Actn44334 Human gen
30	505.8	79.7	10499	3 ABN97929	Abn97929 Human ret
31	500.8	78.9	285020	11 ACN44958	Actn44958 Human gen
32	499.6	78.7	1894	4 AAI14608	Aai14608 Probe #45
33	499.6	78.7	1894	4 ABA56337	Abas6337 Human foe
34	499.6	78.7	1894	4 AAI35980	Aai35980 Probe #46
35	499.6	78.7	1894	4 ABA45822	Abas45822 Human bre
36	499.6	78.7	1894	4 ABA25978	Abas25978 Probe #44
37	499.6	78.7	1894	4 AAK30018	Aak30018 Human bon
38	499.6	78.7	1894	4 AAK04516	Aak04516 Human bra
39	499.6	78.7	1894	4 ABS29670	Abs29670 Human liv
40	499.6	78.7	1894	5 AAI04422	Aai04422 Probe #44
41	499.6	78.7	1894	6 ABS04589	Abs04589 Human gen
42	497.2	78.3	2411	12 ADQ67414	Adq67414 Novel hum
43	495.6	78.0	3372	2 AAX25663	Aax25663 Human end
44	495.6	78.0	3372	3 AAA59213	Aaa59213 Partial p
45	493.2	77.7	2052	4 ABA08902	Abas08902 Human rec

ALIGNMENTS

RESULT 1
AAV43215
ID AAV43215 standard; cDNA; 635 BP.
XX
AC AAV43215;
XX
DT 29-DEC-1998 (first entry)
XX
DE Multiple sclerosis associated retrovirus fragment 4.
XX
KW Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW env gene; rheumatoid arthritis-associated virus; ss.
XX
OS Multiple sclerosis associated retrovirus.
FH Key Location/Qualifiers
CDS 1..234
/*tag= a
/product= "Encodes protein AAV71067"

WO9823755-A1.
XX
XX 04-JUN-1998.
XX
XX 26-NOV-1997; 97MO-IB001482.
XX
XX 26-NOV-1996; 96US-00756429.
XX (INMR) BIO MERIEUX.
XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
XX Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
XX WPI; 1998-322732/28.
XX P-PSDB; AAW71067.
XX
XX New nucleic acid from retroviruses - useful for diagnosis, prevention and
XX treatment of, e.g. multiple sclerosis.
XX
XX Disclosure; Page 183; 286pp; English.
XX
XX The present sequence represents a multiple sclerosis (MS) associated
XX retrovirus (MSRV) genomic fragment used in the method of the invention.

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

CC The invention provides complete or partial genomic sequences of the MSRV-
CC 1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC genes. The invention also provides antibodies raised against the
CC polypeptides. The genomic sequences, polypeptides and antibodies are also
CC claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC associated viruses, and also for prevention and treatment of infection
CC with these viruses
XX
SQ Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;

Query Match 100.0%; Score 635; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 1e-186;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
Db 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGGACCC 120
Db 61 CAAATTTGTTCTTCAAAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGGACCC 120
QY 121 CTGGACGGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
Db 121 CTGGACGGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
Db 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTGTAGAGGGGGGACTGAGAGAC 300
Db 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTGTAGAGGGGGGACTGAGAGAC 300
QY 301 AGACTAGCTGGATTTCTTAGGCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 301 AGACTAGCTGGATTTCTTAGGCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
QY 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTCAACCCGACCAATCAGAGAGCTC 420
QY 421 ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCCTG 480
Db 421 ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCCTG 480
QY 481 AGAGCAGCGGGAGGACAGGATCGGGATATAAACCAGGCATTCGAGCGGCAACGG 540
Db 481 AGAGCAGCGGGAGGACAGGATCGGGATATAAACCAGGCATTCGAGCGGCAACGG 540
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
Db 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600

RESULT 2

AA29702
ID AAX29702 standard; DNA; 635 BP.
XX
AC AAX29702;
XX
DT 17-OCT-2003 (revised)
DT 27-AUG-2003 (revised)
DT 08-JUN-1999 (first entry)
XX
DE Clone CL6-3' from MSRV-1.
XX Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW rheumatoid polyarthritis; ss.

XX Viruses.
XX OS
XX PN FR2765588-A1.
XX
XX
PD 08-JAN-1999.
XX
XX PF 07-JUL-1997; 97FR-00008816.
XX
XX PR 07-JUL-1997; 97FR-00008816.
XX (INMR) BIO MERIEUX.
XX
XX WPI; 1999-098275/09.
XX P-PSDB; AAW9553.
XX
XX Nucleic acid sequences of retrovirus called MSRV-1 - associated with
XX multiple sclerosis or rheumatoid polyarthritis.
XX
XX Claim 1; Page 36-37; 83pp; French.
XX
XX This sequence represents clone CL6-3' from a novel multiple sclerosis
XX related virus type 1 (MSRV1). The sequence can be used in diagnostic,
XX prophylactic or therapeutic compositions to inhibit expression of a
XX multiple sclerosis related virus and/or virus associated with rheumatoid
XX polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
XX 17-OCT-2003 to standardise OS field)
XX
XX Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 635; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 1e-186;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
Db 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGGACCC 120
Db 61 CAAATTTGTTCTTCAAAATGAGACACAGATGGAGTGCATGACTAAGATCCACCGTGGACCC 120
QY 121 CTGGACGGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
Db 121 CTGGACGGGCTCTAGCCATCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
Db 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCACGGGGAAGCATTTAGAGCGGT 240
QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTGTAGAGGGGGGACTGAGAGAC 300
Db 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTGTAGAGGGGGGACTGAGAGAC 300
QY 301 AGGACTAGCTGGATTTCTTAGGCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 301 AGGACTAGCTGGATTTCTTAGGCCAACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
QY 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 361 GCATCCACCTCTAAACATGSGGCTTGCACTTATAGCTCAACCCGACCAATCAGAGAGCTC 420
QY 421 ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCCTG 480
Db 421 ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCCTG 480
QY 481 AGAGCAGCGGGAGGACAGGATCGGGATATAAACCAGGCATTCGAGCGGCAACGG 540
Db 481 AGAGCAGCGGGAGGACAGGATCGGGATATAAACCAGGCATTCGAGCGGCAACGG 540
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
Db 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCTCT 600

QY	601	ATTAATCTTGCACCTGAAATGAAAAA	635
DB	601	ATTAATCTTGCACCTGAAATGAAAAA	635
RESULT 3			
ADG14843	ADG14843 standard; cDNA; 635 BP.		
XX	AC		
XX	AC		
XX	AC		
DT	26-FEB-2004	(first entry)	
XX	MSRV associated cDNA #7.		
DE	es; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.		
XX	Multiple sclerosis associated retrovirus.		
OS	US2003198647-A1.		
XX	23-OCT-2003.		
XX	03-APR-2002; 2002US-00114104.		
XX	26-NOV-1996; 96US-00756429.		
PR	26-NOV-1997; 97US-00979847.		
XX	(INMR) BIO MERIEUX.		
XX	Perron H, Beseme F, Bedin F, Paranhos-Baccala G;		
PI	Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;		
XX	WPI; 2004-032461/03.		
XX	New isolated nucleic acid and their fragments having the pol gene of a		
PT	retrovirus, useful for diagnosing, preventing and/or treating multiple		
PT	sclerosis and/or rheumatoid arthritis.		
XX	Disclosure; SEQ ID NO 102; 193pp; English.		
XX	The invention relates to an isolated nucleic acid which comprises the pol		
CC	gene of a retrovirus associated with multiple sclerosis or rheumatoid		
CC	arthritis. The methods and compositions of the present invention are		
CC	useful for diagnosing, preventing and/or treating multiple sclerosis		
CC	and/or rheumatoid arthritis. The present sequence is used in the		
CC	exemplification of the invention.		
XX	Sequence 635 BP; 188 A; 170 C; 136 G; 141 T; 0 U; 0 Other;		
SQ	Query Match 100.0%; Score 635; DB 12; Length 635;		
	Best Local Similarity 100.0%; Pred. No. 1e-186;		
	Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	CCCTGTATCTTTAACTTCTTAACTTCTTCTCCAGATCAAACTGTAARACTA	60
DB	1	CCCTGTATCTTTAACTTCTTAACTTCTTCTCCAGATCAAACTGTAARACTA	60
QY	61	CAAAATCTTCTTCAATGGAGCACCAGATGGAGTCCATGATCAAGATCCACCGTGACCC	120
DB	61	CAAAATCTTCTTCAATGGAGCACCAGATGGAGTCCATGATCAAGATCCACCGTGACCC	120
QY	121	CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGCAGCCCTCCCGAG	180
DB	121	CTGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGCAGCCCTCCCGAG	180
QY	181	GAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT	240
DB	181	GAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT	240
QY	241	CATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGAGCTGAGAGAC	300
DB	241	CATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGAGCTGAGAGAC	300
DB	241	CATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGAGCTGAGAGAC	300
QY	301	AGGACTAGCTGATTTCTTAGGCCAAACGAAAGATCCCTAAGCTAGCTGGGAGGTGACT	360
DB	301	AGGACTAGCTGATTTCTTAGGCCAAACGAAAGATCCCTAAGCTAGCTGGGAGGTGACT	360
QY	361	GCATCCACCTCTAAACATGCGGCTTGCACCTTAGCTCAACCCGACCAATTCAGAGAGCTC	420
DB	361	GCATCCACCTCTAAACATGCGGCTTGCACCTTAGCTCAACCCGACCAATTCAGAGAGCTC	420
QY	421	ACTAAATCTTAATTTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTCTG	480
DB	421	ACTAAATCTTAATTTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTCTG	480
QY	481	AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCCGCAACGG	540
DB	481	AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCCGCAACGG	540
QY	541	CAACCCCTTTGGTCCCTTCTTGTATGGCGCTCTGTTTCACTCTATTTCCTCTCT	600
DB	541	CAACCCCTTTGGTCCCTTCTTGTATGGCGCTCTGTTTCACTCTATTTCCTCTCT	600
QY	601	ATTAATCTTGCACCTGAAATGAAAAA	635
DB	601	ATTAATCTTGCACCTGAAATGAAAAA	635
RESULT 4			
ADB84397	ADB84397 standard; DNA; 635 BP.		
XX	AC		
XX	AC		
DT	04-DEC-2003	(first entry)	
XX	MSRV-1 associated DNA sequence #17.		
XX	db; multiple sclerosis; rheumatoid arthritis; gag; pol;		
KW	reverse transcriptase; ribonuclease H.		
XX	Unidentified;.		
XX	US2003039664-A1.		
XX	27-FEB-2003.		
XX	26-NOV-1997; 97US-00979847.		
XX	26-NOV-1996; 96US-00756429.		
XX	(PERR/) PERRON H.		
PA	(BESE/) BESEME F.		
PA	(BEDI/) BEDIN F.		
PA	(PARA/) PARANHOS-BACCALA G.		
PA	(KOMU/) KOMURIAN-PRADEL F.		
PA	(JOLI/) JOLIVET-REYNAUD C.		
PA	(MAND/) MANDRAND B.		
PA	(GARS/) GARSON J A.		
PA	(TUKE/) TUKE P W.		
XX	Perron H, Beseme F, Bedin F, Paranhos-Baccala G;		
PI	Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;		
XX	WPI; 2003-512253/48.		
DR	New isolated or purified nucleic acid associated with multiple sclerosis		
XX	and/or rheumatoid arthritis, useful for detecting a virus associated with		
PT	multiple sclerosis or rheumatoid arthritis in a biological sample.		
XX	Claim 31; Page 76-77; 193pp; English.		
XX	The invention relates to an isolated or purified nucleic acid from a		
CC	virus associated with multiple sclerosis and/or rheumatoid arthritis,		
CC			

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs). Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is a claimed MSRV-associated sequence whose identity cannot be
CC accurately determined. Note: The SEQ ID numbers for the sequences as
CC displayed in the main body of the patent do not match the SEQ ID numbers
CC in the sequence listing. Consequently those sequences mentioned in the
CC claims may not be the sequences the authors intended to claim.

XX Sequence 635 BP; 188 A; 167 C; 133 G; 138 T; 0 U; 9 Other;

Query Match 98.6%; Score 626; DB 9; Length 635;
Best Local Similarity 98.6%; Pred. No. 6.4e-184;
Matches 626; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTTCCAGAACTCAAACTCTTAAACTA 60
DB 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTTCCAGAACTCAAACTCTTAAACTA 60
QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB 61 CAAATTTGTTCTTCAAAATGAGACACAGATGGAGTCCATGACTAAGATCCACCGTGACCN 120
QY 121 CTGGACGGGCTGTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 180
DB 121 CTGGACGGGCTGTAGCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAN 180
QY 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGACGGT 240
DB 181 GAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGACGGN 240
QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTGAGAGGGGGACTGAGAGAC 300
DB 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTGAGAGGGGGACTGAGAGAN 300
QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAATCCCTAAGCTAGCTGGGAGGTGACT 360
DB 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAATCCCTAAGCTAGCTGGGAGGTGACN 360
QY 361 GCATCCACCTCTAAACATGGGCTTCAACTAGCTCACACCCGACCAATCAGAGAGCTC 420
DB 361 GCATCCACCTCTAAACATGGGCTTCAACTAGCTCACACCCGACCAATCAGAGAGCTN 420
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTTAAAGAAATAGCAATCATCTATTGCTTG 480
DB 421 ACTAAATGCTAATAGGCAAAATAGGAGTTAAAGAAATAGCAATCATCTATTGCTTN 480
QY 481 AGAGCAGCGGAGGACAGAGATCGGATATAAACCCAGGATTCGAGCGGCAACGG 540
DB 481 AGAGCAGCGGAGGACAGAGATCGGATATAAACCCAGGATTCGAGCGGCAACGN 540
QY 541 CAACCCCTTTGGTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 541 CAACCCCTTTGGTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCN 600
QY 601 ATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 635
DB 601 ATTAATCTTGGCACTGAAAAAAGAAAAAAGAAAAA 635

RESULT 5
AAA63826

ID AAA63826 standard; DNA; 2030 BP.

XX AAA63826;

XX 06-AUG-2003 (revised)

DT 04-DEC-2000 (first entry)

XX Nucleotide sequence of the MSRV-1 3' env and LTR regions.

XX MSRV-1; pol region; long terminal repeat; LTR; RUS region; retrovirus;
KW ss.

XX Multiple sclerosis associated retrovirus.

XX Key Location/Qualifiers

FT CDS 1..1629

FT /note= "Contains one termination codon"

FT sig_peptide 1..81

FT CAAT_signal 1800..1807

FT CAAT_signal 1858..1864

FT TATA_signal 1906..1911

FT polyA_signal 1996..2002

XX WO200047745-A1.

XX 17-AUG-2000.

XX 15-FEB-2000; 2000WO-IB000159.

XX 15-FEB-1999; 99EP-00420041.

XX (INNR) BIO MERIEUX.

XX Paranhos-Baccala G, Perron H, Komurian-Pradel F;

XX WPI; 2000-506097/45.

XX P-PSDB; AAB08195.

XX Nucleotide fragment of LTR-RUS region from Multiple Sclerosis retrovirus

XX (MSRV) used to detect the presence of MSRV-1 retrovirus in a biological

XX sample.

XX Disclosure; Fig 2; 23pp; English.

XX The present sequence represents the nucleotide sequence corresponding to
the 3' env region and long terminal repeat sequences from clone CL6 of
Multiple Sclerosis retrovirus (MSRV-1). The specification describes a
long terminal repeat (LTR)-RUS region which encodes the expression of a
MSRV-1 protein. This is unusual for LTRs, in particular in the RUS
region. The sequence includes CAAT and TATA signals which are present in
the U3 and R regions and are not directed towards the CDS indicated in
the features table. Probes and antibodies to the MSRV-1 retrovirus
protein and encoding polynucleotide sequences are used to detect the
presence of MSRV-1 retrovirus in a biological sample. (Updated on 06-AUG-
2003 to correct OS field.)

XX Sequence 2030 BP; 574 A; 559 C; 387 G; 510 T; 0 U; 0 Other;

Query Match 97.2%; Score 617.4; DB 3; Length 2030;

Best Local Similarity 98.3%; Pred. No. 5.4e-181;

Matches 624; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTCTTGAAGTTTGTCTTCCAGAACTCAAAACTGTAATACTA 60

Db 1396 CCCTGTATCTTTAACTCTCTGTTAAAGTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 1455
 Qy 61 CAATTTGTTCTTCAATAGGACACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 1456 CAGATGGTCTTACAAATAGGAAACCCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 1515
 Qy 121 CTGGACCGGCTGCTAGCCCATCTCCGATGTTAAATGACATGAAGGACCCCTCCCGAG 180
 Db 1516 CTGGACCGGCTGCTAGCCCATCTCCGATGTTAAATGACATGAAGGACCCCTCCCGAG 1575
 Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTTAGAGCGGT 240
 Db 1576 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTTAGAGCGGT 1635
 Qy 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 300
 Db 1636 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 1695
 Qy 301 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 1696 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 1755
 Qy 361 GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCGACCAATTCAGAGAGCTC 420
 Db 1756 GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCGACCAATTCAGAGAGCTC 1815
 Qy 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 480
 Db 1816 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 1875
 Qy 481 AGAGCAGCGGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCCCGGCAACGG 540
 Db 1876 AGAGCAGCGGGAGGACAGGATCGGATATAAACCCAGGATTCGAGCCCGGCAACGG 1935
 Qy 541 CAACCCCTCTTGGTCCCTCCCTTTGTATGGCGCTCTCTTTTCACTCTATTTCACCTCT 600
 Db 1936 CAACCCCTCTTGGTCCCTCCCTTTGTATGGCGCTCTCTTTTCACTCTATTTCACCTCT 1995
 Qy 601 ATTAATCTTGCACCTGAAAAAAGAAAAAAGAAAAA 635
 Db 1996 ATTAATCTTGCACCTGAAAAAAGAAAAAAGAAAAA 2030

RESULT 6

AAD41225
 ID AAD41225 standard; cDNA; 2074 BP.

XX AAD41225;

AC 30-OCT-2002 (first entry)

XX Human EMBRY-2 cDNA.

Human; embryogenesis associated protein; AIDS; reproductive disorder; infertility; endometriosis; endometrial tumour; inflammatory disorder; autoimmune disorder; acquired immune deficiency syndrome; transgenic; ovarian tumour; contact dermatitis; placenta disorder; preclampsia; EMBRY-2; allergy; gene therapy; Gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 64..1692

FT /*tag= a

FT /product= "EMBRY-2 protein"

FT sig_peptide 64..123

FT /*tag= b

FT mat_peptide 124..1689

FT /*tag= c

FT /product= "Mature EMBRY-2 protein"

PN W0200248362-A2.

XX 20-JUN-2002.
 XX 14-NOV-2001; 2001WO-US043956.
 XX 15-NOV-2000; 2000US-0249407P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Ramkumar J, Arvizu C;
 XX WPI; 2002-537629/57.
 XX P-PSDB; A825054.
 XX New polypeptides of human embryogenesis associated proteins for screening
 PT modulators useful for treating or preventing disorders e.g.
 PT endometriosis, infertility, allergy, preclampsia.
 XX Claim 59; Page 96-97; 97pp; English.
 XX The invention relates to human embryogenesis associated proteins (EMBRV)
 CC and nucleic acid molecules encoding such proteins. EMBRY sequences are
 CC useful for screening modulators useful for treating or preventing
 CC disorders associated with abnormal expression of EMBRY. The disorders
 CC treated include reproductive disorders such as infertility.
 CC endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
 CC disorder such as acquired immune deficiency syndrome (AIDS), allergies,
 CC contact dermatitis; disorders of the placenta such as preclampsia,
 CC abruptio placentae etc. Sequences of the invention are also useful for
 CC analysing a proteome of a tissue or a cell type. EMBRY proteins are
 CC useful as immunogens for preparing antibodies. Polynucleotides of the
 CC invention are useful for creating knockin humanised animals or transgenic
 CC animals to model human diseases. They are also used in gene therapy. The
 CC present sequence is human EMBRY-2 cDNA
 XX Sequence 2074 BP; 583 A; 567 C; 392 G; 531 T; 0 U; 1 Other;

Query Match 86.0%; Score 546.4; DB 6; Length 2074;

Best Local Similarity 93.8%; Pred. NO. 6.7e-159;

Matches 579; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

Qy 1 CCCTGTATCTTTAACTCTCTGTTAAAGTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 60
 Db 1459 CCCTGTATCTTTAACTCTCTGTTAAAGTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 1518
 Qy 61 CAATTTGTTCTTCAATAGGACACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 1519 CAATTTGTTCTTCAATAGGACACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 1578
 Qy 121 CTGGACCGGCTGCTAGCCCATCTCCGATGTTAAATGACATGAAGGACCCCTCCCGAG 180
 Db 1579 CTGGACCGGCTGCTAGCCCATCTCCGATGTTAAATGACATGAAGGACCCCTCCCGAG 1638
 Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTTAGAGCGGT 240
 Db 1639 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCATTTAGAGCGGT 1698
 Qy 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 300
 Db 1699 CGTCAGTCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 1758
 Qy 301 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 1759 AGGACTAGCTGGATTTCCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 1818
 Qy 361 GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCGACCAATTCAGAGAGCTC 420
 Db 1819 GCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCGACCAATTCAGAGAGCTC 1878
 Qy 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 480
 Db 1879 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCCTG 1938

[illegible]

RESULT 7	
AAV43219	
ID	AAV43219 standard; CDNA; 1329 BP.
XX	
AC	AAV43219;
XX	
DT	29-DEC-1998 (first entry)
XX	
DE	Multiple sclerosis associated retrovirus fragment 6.
XX	
KW	Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KM	env gene; rheumatoid arthritis-associated virus; ss.
XX	
OS	Multiple sclerosis associated retrovirus.
XX	
Key	Location/Qualifiers
CDS	2..490
FT	/*tag= a
FT	/product= "Encodes protein AAV71069"
FT	/transl_except= (pos:77-79, appears to code for a stop
FT	codon)
FT	/transl_except= (pos:125-127, appears to code for a stop
FT	codon)
FT	/transl_except= (pos:137-139, appears to code for a stop
FT	codon)

XX	W09823755-A1.
XX	PN
XX	04-JUN-1998.
XX	PD
XX	PF
XX	26-NOV-1997; 97WO-IB001482.
XX	XX
XX	26-NOV-1996; 96US-00756429.
XX	PR
XX	(INNR) BIO MERIEUX.
XX	PA
XX	Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI	Komurian-Pradel F, Jollivet-Reynaud C, Mandrand B;
PI	XX
XX	WPI; 1998-322732/28.
DR	DR
DR	P-PSDB; AAW71069.
XX	XX
PT	New nucleic acid from retroviruses - useful for diagnosis, prevention and treatment of, e.g. multiple sclerosis.
PT	XX
PS	Disclosure; Page 187-188; 286pp; English.
XX	XX
CC	The present sequence represents a multiple sclerosis (MS) associated
CC	retrovirus (MSRV) genomic fragment used in the method of the invention.
CC	CC The invention provides complete or partial genomic sequences of the MSRV-
CC	1 pol gene, gag gene and env gene, and polypeptides encoded by these
CC	genes. The invention also provides antibodies raised against the
CC	polypeptides. The genomic sequences, polypeptides and antibodies are also
CC	claimed useful for diagnosing infection by MS and rheumatoid arthritis-
CC	associated viruses, and also for prevention and treatment of infection
CC	with these viruses
XX	XX
XX	Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;

	Best Local Similarity	92.4%;	Pred. No. 1.5e-158;			
	Matches	572;	Conservative	0;	Mismatches	47;
					Indels	0;
					Gaps	0;
QY	1	CCCTGTATCTTTAAACCTCTTGTAAAGTTTGTCTCTTCCAGAATCAAAACTGTAAACTA	60			
Db	257	CCCTGTATCTCAACTTCTCTGTTAAAGTTTGTCTCTTCCAGAATTGAAGCTGTAAAGCTA	316			
QY	61	CAAAATGTTCTTCAAAATGGAGCAGACAGATGGATGCTACATGAAGATCCACCGTGACCC	120			
Db	317	CAAAATGTTCTTCAAAATGGAAACCCAGATGCACTCATGACTAAATCTACCGTGACCC	376			
QY	121	CTGGACCGGCTGTAGCCCATGCTCCGATGTTAAATGACATTTGAAGGACCCCTCCCGAG	180			
Db	377	CTGGACCGGCTGTAGACTATGCTCTGATGTTAAATGACATTTGAAGTCAACCCCTCCCGAG	436			
QY	181	GAATCTCTCACTGACCAACCCCTACTATGCCCAATTCACGGGAAGCAGTTAGACGGT	240			
Db	437	GAATCTCTCACTGACCAACCCCTACTACTCTCAATTCAGTAGGAAGCAGTTAGACAGT	496			
QY	241	CATCAGCCAACTCCGCCAAACAGCAGCTTGGGTTTCTGTGTAGAGGGGGACTTGAGAGAC	300			
Db	497	TGTCAGCCAACTCCGCCAAACAGTACTTGGGTTTCTGTGTAGAGGGTGGACTTGAGAGAC	556			
QY	301	AGGACTAGCTGATTTCTTAGGCCACGAGAAATCCCTAAGCCTAGCTGGGAAGGTTGACT	360			
Db	557	AGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCAAGCCCTANCTGGGAAGTGAAC	616			
QY	361	GCATCCACCTCTAAACATGGGGCTTGCACATTTAGCTCACACCGACCAATCAGAGAGCTC	420			
Db	617	GCATCCATCTTTAAACATGGGGCTTGCACATTTAGCTCACACCGACCAATCAGAGAGCTC	676			
QY	421	ACTAAATGCTTAATTAGGCNAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCGTG	480			
Db	677	ACTAAATGCTTAATCAGGCNAAACAGGAGGTAAAGCAATAGCCAATCATCTATTGCGTG	736			
QY	481	AGAGCAGCGGGAGGAGCAAGGATCGGGAATTAACCCAGGCAATTCGAGCCGCCACGG	540			
Db	737	AGAGCAGCGGGAGGAGCAAGGATCGGGAATTAACCTCAGGCAATTCAGGCCAGCAACAG	796			
QY	541	CAACCCGCTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTATTTCACTCT	600			
Db	797	CAACCCGCTTGGGTCCCTCCCATTTGTATGGGAGCTCTGTTTCACTCTATTATTTCACTCT	856			
QY	601	ATTAAATCTTGCACCTGAA	619			
Db	857	ATTAAATCTGCAACTGCA	875			

RESULT 8	
AAAX29704	
ID	AAAX29704 standard; DNA; 1329 BP.
XX	
XX	
AC	AAAX29704;
XX	
XX	17-OCT-2003 (revised)
DT	27-AUG-2003 (revised)
DT	08-JUN-1999 (first entry)
DT	
XX	
XX	Clone 5M6 from MSRV-1.
DE	
XX	
XX	Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
KW	rheumatoid polyarthritis; ss.
KW	
XX	
XX	Viruses.
OS	
XX	
XX	FR2765588-A1.
PN	
XX	
PD	08-JAN-1999.
XX	
XX	07-JUL-1997; 97FR-00008816.
PF	
XX	
PR	07-JUL-1997; 97FR-00008816.
XX	
XX	

PA (INMR) BIO MERIEUX.
XX WPI; 1999-098275/09.
DR P-PSDB; AAW99554.
XX Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT multiple sclerosis or rheumatoid polyarthritis.
XX Claim 1; Page 39-40; 83pp; French.
XX This sequence represents clone 5M6 from a novel multiple sclerosis
CC related virus type 1 (MSRV1). The sequence can be used in diagnostic,
CC prophylactic or therapeutic compositions to inhibit expression of a
CC multiple sclerosis related virus and/or virus associated with rheumatoid
CC polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
CC 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;
Query Match 85.8%; Score 545; DB 2; Length 1329;
Best Local Similarity 92.4%; Pred. No. 1.5e-158;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 1 CCCTGTATCTTTAACTCTCTTGTAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
Db 257 CCCTGTATCTTCACTCTCTTGTAGTTTGTCTTCCAGAAATGAAGCTGTAAAGCTA 316
QY 61 CAAATTTGTTCTTCAAATGGAGCCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 317 CAAATAGTTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAAATCTACCGTGACCC 376
QY 121 CTGACCGGCTGCTAGCCATGCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 180
Db 377 CTGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 436
QY 181 GAAATCTCAATCTGACACACCCCTACTATGCCCAATTCAGCGGGAACAGTGTAGAGCGT 240
Db 437 GAAATCTCAATCTGACACACCCCTACTACTACCTCCATTCAGTAGAAGCAGTGTAGAGCAGT 496
QY 241 CATCAGCCAACTCTCCCAACAGCACTTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGC 300
Db 497 TGTGACCGGCTGCTAGCCATGCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 556
QY 301 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 360
Db 557 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 436
QY 181 GAAATCTCAATCTGACACACCCCTACTATGCCCAATTCAGCGGGAACAGTGTAGAGCGT 240
Db 437 GAAATCTCAATCTGACACACCCCTACTACTACCTCCATTCAGTAGAAGCAGTGTAGAGCAGT 496
QY 241 CATCAGCCAACTCTCCCAACAGCACTTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGC 300
Db 497 TGTGACCGGCTGCTAGCCATGCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 556
QY 301 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 360
Db 557 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 436
QY 361 GCATCCACCTCTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 617 GCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 676
QY 421 ACTAAATGCTAAATAGGCAAAAATAGAGGTAAAGAAATAGCCAATCATCTATTGCGTG 480
Db 677 ACTAAATGCTAAATAGGCAAAAATAGAGGTAAAGCAATAGCCAATCATCTATTGCGTG 736
QY 481 AGAGCAGCGGGAGGACAGAGATCGGATATATAACCCAGGCAATTCGAGCCGCAACGG 540
Db 737 AGAGCAGCGGGAGGACAGAGATCGGATATATAACCCAGGCAATTCGAGCCGCAACAG 796
QY 541 CAACCCCTTTGGGCTCCCTCCCTTTGTTATGGGGCTCTGTTTCACTCTATTTCCTCT 600
Db 797 CAACCCCTTTGGGCTCCCTCCCTTTGTTATGGGGAGCTCTGTTTCACTCTATTTCCTCT 856
QY 601 ATTAATCTTGCACACTGAA 619
Db 857 ATTAATCATGCAACTGCA 875
RESULT 9
ID ADG14849
XX ADG14849 standard; cDNA; 1329 BP.
AC ADG14849;

XX 26-FEB-2004 (first entry)
XX MSRV associated cDNA #11.
XX ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX Multiple sclerosis associated retrovirus.
XX US2003198647-A1.
XX 23-OCT-2003.
XX 03-APR-2002; 2002US-00114104.
XX 26-NOV-1996; 96US-00756429.
XX 26-NOV-1997; 97US-00979847.
XX (INMR) BIO MERIEUX.
XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX WPI; 2004-032461/03.
XX New isolated nucleic acid and their fragments having the pol gene of a
PT retrovirus, useful for diagnosing, preventing and/or treating multiple
PT sclerosis and/or rheumatoid arthritis.
XX Disclosure; SEQ ID NO 108; 193pp; English.
XX The invention relates to an isolated nucleic acid which comprises the pol
CC gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC arthritis. The methods and compositions of the present invention are
CC useful for diagnosing, preventing and/or treating multiple sclerosis
CC and/or rheumatoid arthritis. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;
Query Match 85.8%; Score 545; DB 12; Length 1329;
Best Local Similarity 92.4%; Pred. No. 1.5e-158;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 1 CCCTGTATCTTTAACTCTCTTGTAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
Db 257 CCCTGTATCTTCACTCTCTTGTAGTTTGTCTTCCAGAAATGAAGCTGTAAAGCTA 316
QY 61 CAAATTTGTTCTTCAAATGGAGCCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 317 CAAATAGTTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAAATCTACCGTGACCC 376
QY 121 CTGACCGGCTGCTAGCCATGCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 180
Db 377 CTGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 436
QY 181 GAAATCTCAATCTGACACACCCCTACTATGCCCAATTCAGCGGGAACAGTGTAGAGCGT 240
Db 437 GAAATCTCAATCTGACACACCCCTACTACTACCTCCATTCAGTAGAAGCAGTGTAGAGCAGT 496
QY 241 CATCAGCCAACTCTCCCAACAGCACTTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGC 300
Db 497 TGTGACCGGCTGCTAGCCATGCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 556
QY 301 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 360
Db 557 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 436
QY 181 GAAATCTCAATCTGACACACCCCTACTATGCCCAATTCAGCGGGAACAGTGTAGAGCGT 240
Db 437 GAAATCTCAATCTGACACACCCCTACTACTACCTCCATTCAGTAGAAGCAGTGTAGAGCAGT 496
QY 241 CATCAGCCAACTCTCCCAACAGCACTTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGC 300
Db 497 TGTGACCGGCTGCTAGCCATGCTCCGATGTTTAATGACATTAAGGACCCCTCCCGAG 556
QY 301 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 360
Db 557 AGGACTAGCTGGATTTCTAGGCTGCTGATGTTTAATGACATTAAGGACCCCTCCCGAG 436
QY 361 GCATCCACCTCTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 617 GCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 676
QY 421 ACTAAATGCTAAATAGGCAAAAATAGAGGTAAAGAAATAGCCAATCATCTATTGCGTG 480
Db 677 ACTAAATGCTAAATAGGCAAAAATAGAGGTAAAGCAATAGCCAATCATCTATTGCGTG 736
QY 481 AGAGCAGCGGGAGGACAGAGATCGGATATATAACCCAGGCAATTCGAGCCGCAACGG 540
Db 737 AGAGCAGCGGGAGGACAGAGATCGGATATATAACCCAGGCAATTCGAGCCGCAACAG 796
QY 541 CAACCCCTTTGGGCTCCCTCCCTTTGTTATGGGGCTCTGTTTCACTCTATTTCCTCT 600
Db 797 CAACCCCTTTGGGCTCCCTCCCTTTGTTATGGGGAGCTCTGTTTCACTCTATTTCCTCT 856
QY 601 ATTAATCTTGCACACTGAA 619
Db 857 ATTAATCATGCAACTGCA 875
RESULT 9
ID ADG14849
XX ADG14849 standard; cDNA; 1329 BP.
AC ADG14849;

Db 677 ACTAAATCTAATCAGGCAAAAACAGGAGTAAAGCAATACCAATCATCTATTGCCTG 736
 QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGATTCGAGCGCGCAACGG 540
 Db 737 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGATTCGAGCGCGCAACAG 796
 QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGGCTCTGTTTCACTCTATTCACTCT 600
 Db 797 CAACCCCTTTGGTCCCTCCCTTTGATGGGCTCTGTTTCACTCTATTCACTCT 856
 QY 601 ATTAATCTTGAACCTGAA 619
 Db 857 ATTAATCATGCAACTGCA 875

RESULT 10
 ACN44454
 ID ACN44454 standard; DNA; 21646 BP.

AC ACN44454;
 XX
 XX 18-NOV-2004 (first entry)
 XX
 XX Human genomic sequence hCG1748307.

XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.
 XX WO2003073826-A2.
 XX

XX 12-SEP-2003.
 XX
 XX 28-FEB-2003; 2003WO-US006235.
 XX
 XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 910; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published

XX Sequence 21646 BP; 5145 A; 5617 C; 5218 G; 5646 T; 0 U; 20 Other;

Query Match 85.6%; Score 543.4; DB 11; Length 21646;
 Best Local Similarity 94.8%; Pred. No. 1.8e-157;
 Matches 562; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCCTCTGTTAAAGTTGCTCTTCCAGATCAAACTGTAAACTA 60
 Db 19571 CCCTGTATCTTTAACTCCTCTGTTAAAGTTGCTCTTCCAGATTTGAAACTGTAAACTA 19630

QY 61 CAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 19631 CAAATGTTCTTCAAAATGGAGCCCCAGATGCAGTCCATGACTAAGATCCACACAGACCC 19690
 QY 121 CTGACCGCGCTGCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCGAG 180
 Db 19691 CTGACCGAGCTGCTAGCCCATGCTCCGATGTTAATGACATCGAAGGACCCCCCGAG 19750
 QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
 Db 19751 GAAATCTCAACTGTCAACCCCTACTATGCCCAATTCAGCGGAGCAGTTACAGCGGT 19810
 QY 241 CATCAGCCAACTCCCAACAGACCTTGGGTTTTCTGTCAGAGGGGGGACTGAGAGAC 300
 Db 19811 CATCAGCCAACTCCCAACAGACCTTGGGTTTTCTGTCAGAGGGGGGACTGAGAGAC 19870
 QY 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 360
 Db 19871 AAGACTAGCTGGATTTCTTAGGCCGACTAAGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 19930
 QY 361 GCATCCACCTCTAAACATGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 420
 Db 19931 GCATCCACCTTTAAACATGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 19990
 QY 421 ACTAAATGCTAAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCTG 480
 Db 19991 ACTAAATGCTAAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCTG 20050
 QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGATTCGAGCGCGCAACGG 540
 Db 20051 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGATTCGAGCGCGCAATGG 20110
 QY 541 CAACCCCTTTGGGTCCTCCCTCTGTTGATGGGCTCTGTTTCACTCTATT 593
 Db 20111 CAACCCCTTTGGGTCCTCCCTCTGTTGATGGGCTCTGTTTCACTCTATT 20163

RESULT 11
 ACC46747

ID ACC46747 standard; cDNA; 2046 BP.

XX AC ACC46747;

XX 02-JUN-2003 (first entry)

XX Human dithp growth/development-associated protein-encoding cDNA.

XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging; growth;
 KW development; gene; ss.

XX Homo sapiens.

XX WO200297031-A2.

XX 05-DEC-2002.

XX 27-MAR-2002; 2002WO-US010056.

XX 28-MAR-2001; 2001US-0279619P.

XX 29-MAR-2001; 2001US-0280067P.

XX 29-MAR-2001; 2001US-0280069P.

XX 16-MAY-2001; 2001US-0291280P.

XX 17-MAY-2001; 2001US-0291829P.

XX 17-MAY-2001; 2001US-0291849P.

XX 19-JUN-2001; 2001US-0299428P.

XX 20-JUN-2001; 2001US-0299776P.

PR 20-JUN-2001; 2001US-0300001P.
XX (INCY-) INCYTE GENOMICS INC.
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
XX P-PSDB; ABR41810.
XX Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX Claim 2; SEQ ID NO 668; 591pp; English.
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a dithp cDNA encoding a DITHP protein
CC which is associated with growth and development. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2046 BP; 587 A; 494 C; 391 G; 566 T; 0 U; 8 Other;
Query Match 85.1%; Score 540.2; DB 8; Length 2046;
Best Local Similarity 94.4%; Pred. No. 5.6e-157;
Matches 560; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Qy 1 CCTGTATCTTTAACTCTTTTAACTTTGCTCTCTCCAGATCAAACTGTAACCTA 60
Db 571 CTCTGTATCTTTAACTCTTTTAACTTTGCTCTCTCCAGATTTGAACTGTAACCTA 630
Qy 61 CAATTTCTTTCAATAGGAGCAGCAGATGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 631 CAATGTTCTTTCAATAGGAGCAGCAGATGAGTCCATGACTAAGATCCACCGTGACCC 690
Qy 121 CTGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCAGCCCTCCCGAG 180
Db 691 CTGACCGAGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCAGCCCTCCCGAG 750
Qy 181 GAAATCTCAATGCAACACCTCTACTATGCCCAATTCAGCGGGAAGCAGTAGCGGT 240
Db 751 GAAATCTCAATGTCACACCTCTACTATGCCCAATTCAGCGGGAAGCAGTAGCGGT 810
Qy 241 CATCAGCCAACTCTCCCAACAGCAGACTTGGGTTTCTTCTGTTGAGAGGGGAGCTCAGAGAC 300

Db 811 CATCAGCCAACTCTCCCAACAGCAGACTTGGGTTTCTTCTGTTGAGAGGGCACTGAGAGAC 870
Qy 301 AGGACTAGCTGGATTTCCTAGGCCCAACGAAGAATCCCTAAGCCTAGCTCGGAAGGTGACT 360
Db 871 AAGACTAGCTGGATTTCCTAGGCCCAACGAAGAATCCCTAAGCCTAGCTCGGAAGGTGACT 930
Qy 361 GCATCCACCTCTAAACATCGGGCTTTGCAACTTTAGCTTCACACCGGACCAATCAGAGAGCTC 420
Db 931 GCATCCACCTTTAAACATCGGGCTTTGCCACTTTAGCTTCACACCAACCAATCAGAGAGCTC 990
Qy 421 ACTAAATGCTAATTTAGGCCAAATAAGAGGTTAAAGAAATAGCCCAATCATCTATTCCCTG 480
Db 991 ACTAAATGCTAATTTAGGCCAAATAAGAGGTTAAAGAAATAGCCCAATCATCTATTCCCTG 1050
Qy 481 AGAGCAGCAGCGGAGGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGCGCAACGG 540
Db 1051 AGAGCAGCAGCAGGAGGCAAGGATCGGGATATAAACCCGCGCAATTCGAGCCGCGCAATGG 1110
Qy 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCTACTATT 593
Db 1111 CAACCCCTTTGGGTCCTCCCTCTTGTATGGGAGCTCTGTTTCTACTATT 1163
RESULT 12
ADB84403
ID ADB84403 standard; DNA; 1329 BP.
XX ADB84403;
AC ADB84403;
DT 04-DEC-2003 (first entry)
XX MSRV-1 associated DNA sequence #21.
XX ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
XX reverse transcriptase; ribonuclease H.
XX Unidentified;
XX US2003039664-A1.
XX 27-FEB-2003.
XX 26-NOV-1997; 97US-00979847.
XX 26-NOV-1996; 96US-00756429.
XX (PERR/) PERRON H.
XX (BESE/) BESEME F.
XX (BEDI/) BEDIN F.
XX (PARA/) PARANHOS-BACCALA G.
XX (KOMU/) KOMURIAN-PRADEL F.
XX (JOLI/) JOLIVET-REYNAUD C.
XX (MAND/) MANDRAND B.
XX (GARS/) GARSON J A.
XX (TUXE/) TUXE P W.
XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
PI WPI; 2003-512253/48.
XX New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX Claim 31; Page 80; 193pp; English.
XX The invention relates to an isolated or purified nucleic acid from a
CC virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,

CC antigenic peptides or conserved motifs). Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is a claimed MSRV-associated sequence whose identity cannot be
CC accurately determined. Note: The SEQ ID numbers for the sequences as
CC displayed in the main body of the patent do not match the SEQ ID numbers
CC in the sequence listing. Consequently those sequences mentioned in the
CC claims may not be the sequences the authors intended to claim.
XX
SQ Sequence 1329 BP; 339 A; 346 C; 282 G; 332 T; 0 U; 30 Other;

Query Match 84.3%; Score 535; DB 9; Length 1329;
Best Local Similarity 90.8%; Pred. No. 1.9e-155;
Matches 562; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60
DB |||||||
QY 61 CAAATTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB |||||||
QY 317 CAAATAGTTCTTCAATGGAAACCCAGATGCAGTCCATGACTANAACTACCGTGACCC 376
DB |||||||
QY 121 CTGACCGGCTCTAGCCATGCTCGATGTTTAATGACATTGAAGGCCACCTCCCGAG 180
DB |||||||
QY 377 CTGGACGGGCTCTAGACTATGCTCTGATGTTTAATGACATTGNACTACCCCTCCGAG 436
DB |||||||
QY 181 GAAATCTCAACTGCACAAACCTTACTATGCCCCAATTCAGCGGGAAGCAGTTAGACGCT 240
DB |||||||
QY 437 GAAATCTCAACTGCACAAACCTTACTATGCCCCAATTCAGTAAAGCAGTTAGACGCT 496
DB |||||||
QY 241 CATAGCCAACTCTCCCAACAGCACTTGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
DB |||||||
QY 497 TGTGAGCAACCTCTCCCAACAGCACTTGGTTTCTGTTGAGNGGGTGGACTGAGAGAC 556
DB |||||||
QY 301 AGGACTAGCTGGATTCTCTAGGCCAACGAAAGATCCTTAAGCCTAGCTGGGAAGGTGACT 360
DB |||||||
QY 557 AGGACTAGCTGGATTCTCTAGGCTGACTAAGATCCNAAAGCCNANCTGGGAAGGTGACC 616
DB |||||||
QY 361 GCATCCACCTCTAAACATGGGGTTGCACTTAGCTCACACCGCAATACAGAGAGTTC 420
DB |||||||
QY 617 GCATCCATCTTTAAACATGGGGTTGCACTTAGCTCACACCGCAATACAGAGAGTTC 676
DB |||||||
QY 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
DB |||||||
QY 677 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGCAATAGCNAATCATCTATTGCTG 736
DB |||||||
QY 481 AGACACAGCGGGAGGACAAAGGATCGGATATAAAACCCAGGATTCGAGCCCGCAACGG 540
DB |||||||
QY 737 AGACACAGCGGGAGGACAAAGGATCGGATATAAACTAGGNTTCAAGCCAGCAACAG 796
DB |||||||
QY 541 CAACCCCTTTGGTCCCTCTCTTTGTATGGGGGCTCTGTTTCACTCTATTTCACCT 600
DB |||||||
QY 797 CAACCCCTTTGGTCCCTCTCTTTATGATGGAGCTCTGTTTNCACCTCTATTTCACCT 856
DB |||||||
QY 601 ATTAATCTTGCACCTGAA 619
DB |||||||
QY 857 ATTAATCTTGCACCTGCA 875
DB |||||||

AAX77526

ID AAX77526 standard; cDNA; 2946 BP.

XX AAX77526;

XX 10-AUG-1999 (first entry)

XX Human secreted protein AJ172_2 cDNA.

XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW cell proliferation; cell differentiation; suppressor; tumour inhibition;
KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
KW cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.

OS Homo sapiens.

XX WO9926972-A1.

XX 03-JUN-1999.

XX 17-NOV-1998; 98WO-US024614.

XX 21-NOV-1997; 97US-00976110.

XX 18-MAY-1998; 98US-00080478.

XX 20-OCT-1998; 98US-00175928.

XX (GEMY) GENETICS INST INC.

XX Jacobs K, McCoey JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Merberg D, Treacy M;

XX WPI; 1999-357813/30.

XX P-PSDB; AAY08622.

XX New polynucleotides encoding secreted proteins.

XX Claim 13a; Page 100-101; 142pp; English.

XX This invention describes novel human secreted proteins encoded by
CC polynucleotides isolated from human adult testes, adult brain, adult
CC blood or adult placenta, or murine adult bone marrow or thymus cDNA
CC libraries. The products of the invention are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful for
CC gene therapy

XX Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

Query Match 84.1%; Score 533.8; DB 2; Length 2946;

Best Local Similarity 91.5%; Pred. No. 6.6e-155;

Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60

DB |||||||

QY 61 CAAATTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120

DB |||||||

QY 121 CTGACCGGCTCTAGCCATGCTCCGATGTTTAATGACATTGAAGGCCACCTCCCGAG 180

DB |||||||

QY 2431 CTGGACCGGCTCTAGCCACGATCTGATGTTAATGACATCAAGGCCACCTCTCTGAG 2490

DB |||||||

181 GAAATCTCACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
|||||
2491 GAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 2550
|||||
241 CATCAGCAACCTCCCAACAGCAGCACTTGGGTTTCTCTTTGAGAGGGGGAGCTGAGAGAC 300
|||||
2551 CGTCGGCCAACTCCCAACAGCAGCACTTAGGTTTCTCTTTGAGATGGGGAGCTGAGAGAC 2610
|||||
301 AGACTAGCTGGATTCTCTAGGCAACAGGAATCCCTAAGCCTAGCTAGTGGGAAGGTGACT 360
|||||
2611 AGGACTAGCTGGATTCTCTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGTGGGAAGGTGACC 2670
|||||
361 GCATCCACCTCTAAACATGGGCTTGCACTTAGCTCACAACCCGACCAATCAGAGAGCTC 420
|||||
2671 ACATCCACCTTTAAACAGGGGCTTGCACTTAGCTCACAACCTGACCAATCAGAGAGCTC 2730
|||||
421 ACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTTG 480
|||||
2731 ACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTTG 2790
|||||
481 AGAGCAGCGGAGGAGCAAGGATCGGGATATAACCCAGGCAATTCGAGCCGGCAACGG 540
|||||
2791 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGTCTTCGAGCCGGCAACGG 2850
|||||
541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCTCTCTTTTCACTCTATTTCACCTCT 600
|||||
2851 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 2910
|||||
601 ATTAATCTTTGCAACTGCAAAAAAAGGAGGTAAAGAAATAGCAATCATCTATTTCACCTCT 635
|||||
2911 ATTAATCTTTGCAACTGCAAAAAAAGGAGGTAAAGAAATAGCAATCATCTATTTCACCTCT 2945
|||||

RESULT 14
AAZ59468
ID AAZ59468 standard; cDNA; 2946 BP.
XX AAZ59468;
XX
11-APR-2000 (first entry)
DE Human secreted protein AJ172_2 polynucleotide sequence.
XX
XX Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
XX placental pathology; metastasis inhibition; nutritional activity;
XX immune stimulator; haematopoiesis regulator; tissue growth;
XX tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
XX gene therapy; ss.
XX Homo sapiens.
XX
XX WO9960020-A1.
XX
XX 25-NOV-1999.
XX
XX 17-MAY-1999; 99WO-US010915.
XX
XX 18-MAY-1998; 98US-00080478.
XX 20-OCT-1998; 98US-00175928.
XX
XX (GEMV) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Mi S, Treacy M;
XX
XX WPI; 2000-116311/10.
XX P-PSDB; AAY67313.
XX
XX New polynucleotides encoding secreted cDNA libraries, used to develop
XX products for the diagnosis and treatment of neoplastic disease.
XX
XX Claim 14; Page 107-108; 149pp; English.
PS

XX This is the human secreted protein AJ172_2 nucleotide sequence, obtained
CC from a human adult testes cDNA library. The invention relates to secreted
CC human and murine proteins. The polynucleotides and proteins are predicted
CC to have biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals. Detection of the levels of the proteins can be used for the
CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
CC which modulate the expression or function of the proteins may be used for
CC treating a neoplastic disease and inhibiting metastasis. Other suggested
CC activities include nutritional activity (e.g. in feeds), cytokine and
CC cell proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotide sequences are also stated to be useful for gene therapy
XX

SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

Query Match 84.1%; Score 533.8; DB 3; Length 2946;
Best Local Similarity 91.5%; Pred. No. 6.6e-155;
Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

'QY 1 CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGAAATCAAACTGTAAACTA 60
Db 2323 CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGAAATCAAACTGTAAACTA 2382
|||||
QY 61 CAATTTGTTCTTCAAAATGAGACACAGATGGATCCATGATCAAGTCCACCTGGACCC 120
Db 2383 -----CAAAATGGAGCCCAAGATGCACTCAAGATCAAGTCTTACCCGAGACCC 2430
|||||
QY 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
Db 2431 CTGGACCGGCTCTAGCCCATGCTCTGATGTTTAATGACATCAAGGACACCCCTCCCGAG 2490
|||||
QY 181 GAAATCTCAACTGCACAAACCCCTTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 2491 GAAATCTCAAGTGCACAACTCTTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 2550
|||||
QY 241 CATCAGCAACCTCCCAACAGCAGCACTTGGGTTTCTCTTTGAGAGGGGGAGCTGAGAGAC 300
Db 2551 CGTCGGCCAACTCCCAACAGCAGCACTTAGGTTTCTCTTTGAGATGGGGAGCTGAGAGAC 2610
|||||
QY 301 AGGACTAGCTGGATTCTCTAGGCAACAGGAATCCCTAAGCCTAGCTGGAAGGTGACT 360
Db 2611 AGGACTAGCTGGATTCTCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGAAGGTGACC 2670
|||||
QY 361 GCATCCACCTCTAAACATGGGCTTGCACCTTAGCTCACAACCCGACCAATCAGAGAGCTC 420
Db 2671 ACATCCACCTTTAAACAGCGGGCTTGCACTTAGCTCACAACCTGACCAATCAGAGAGCTC 2730
|||||
QY 421 ACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTTCCTG 480
Db 2731 ACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTTCCTG 2790
|||||
QY 481 AGAGCAGCGGAGGAGCAAGGATCGGGATATAACCCAGGCAATTCGAGCCGGCAACGG 540
Db 2791 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGTCTTCGAGCCGGCAACGG 2850
|||||
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 600
Db 2851 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 2910
|||||
QY 601 ATTAATCTTTGCAACTGCAAAAAAAGGAGGTAAAGAAATAGCAATCATCTATTTCACCTCT 635
Db 2911 ATTAATCTTTGCAACTGCAAAAAAAGGAGGTAAAGAAATAGCAATCATCTATTTCACCTCT 2945
|||||

RESULT 15
ADC38776
ID ADC38776 standard; cDNA; 2946 BP.

XX AC ADC38776;
XX DT 18-DEC-2003 (first entry)
XX DE Human cDNA encoding a secreted protein #63.
XX ss; gene; immune disorder; severe combined immunodeficiency; SCID;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;
KW lymphoid cell deficiency; osteoporosis; osteoarthritis;
KW peripheral nervous system disease; peripheral neuropathy;
KW Alzheimer's disease; Parkinson's disease; coagulation disorder;
KW inflammatory disease; systemic inflammatory response syndrome; SIRS;
KW ischaemia-reperfusion injury; Crohn's disease; anaphylaxis; fertility;
KW hypersensitivity; regeneration; neural cell proliferation; tumour;
tumour; chemokine; human; secreted protein.
XX Homo sapiens.
OS US2002193567-A1.
XX PN 19-DEC-2002.
XX PD 02-APR-2002; 2002US-00114893.
XX PF 11-AUG-1995; 95US-00514014.
XX PR 05-APR-1996; 96US-00628364.
XX PR 19-APR-1996; 96US-00635311.
XX PR 07-JUN-1996; 96US-00659224.
XX PR 17-JUN-1996; 96US-00664596.
XX PR 09-JUL-1996; 96US-00677231.
XX PR 26-JUL-1996; 96US-00686878.
XX PR 23-AUG-1996; 96US-00701819.
XX PR 27-SEP-1996; 96US-00721488.
XX PR 27-SEP-1996; 96US-00721798.
XX PR 27-SEP-1996; 96US-00721923.
XX PR 27-SEP-1996; 96US-00721926.
XX PR 25-OCT-1996; 96US-00738367.
XX PR 30-OCT-1996; 96US-00739775.
XX PR 13-JAN-1997; 97US-00783395.
XX PR 10-APR-1997; 97US-00833823.
XX PR 02-JUN-1997; 97US-00867677.
XX PR 05-SEP-1997; 97US-00924838.
XX PR 06-OCT-1999; 99US-00413232.
XX PA (GEM) GENETICS INST INC.
XX PI Jacobs K, McCoy JM, Lavallie BR, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;
PI Kelleher K;
XX WPI; 2003-657236/62.
XX P-PSDB; ADC38777.
XX Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and
PT BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,
PT useful for treating e.g. multiple sclerosis and rheumatoid arthritis.
XX Disclosure; SEQ ID NO 134; 412pp; English.
XX The invention relates to a protein comprising fully defined AZ302 1
CC protein or BD127 1 6 protein. The polynucleotides are useful for
CC expressing recombinant proteins for analysis and are also useful as
CC chromosome markers or tags to identify chromosomes or to map related gene
CC positions. The proteins are useful as amino acid supplement, carbon
CC source, nitrogen source and carbohydrate source. The proteins are useful
CC for treating various immune deficiencies and disorders (e.g. severe
CC combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic
CC reactions (e.g. asthma), myeloid or lymphoid cell deficiencies,
CC osteoporosis or osteoarthritis, peripheral nervous system disease (e.g.
CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),

CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory
CC response syndrome (SIRS)), ischaemia-reperfusion injury, Crohn's disease),
CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing
CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or
CC nerve growth or regeneration, for proliferating neural cells and for
CC regenerating nerve and brain tissue, for inducing fertility and for
CC inhibiting tumour growth. Proteins are also useful as chemokine for
CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also
CC useful as inhibitors of receptor/ligand interactions. The present
CC sequence represents cDNA encoding a human secreted protein.

XX SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;
Query Match 84.1%; Score 533.8; DB 10; Length 2946;
Best Local Similarity 91.5%; Pred. No. 6.6e-155;
Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAACTCAAACTGTAACCTA 60
DB 2323 CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAACTCAAACTGTAACCTA 2382
QY 61 CAAATTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 2383 -----CAAATGGAGCCCAAGATCGAGTCCAGACTAAGACTACCGCAGACCC 2430
QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAAATGACATTTGAAGGCACCCCTCCCGAG 180
DB 2431 CTGGACCGGCTGCTAGCCCATGCTGATGTTTAAATGACATCAAAAGGCACCCCTCCCGAG 2490
QY 181 GAAATCTCAATGTCACAAACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
DB 2491 GAAATCTCAGCTGCACAAACCTCTACTAGCCCAATTCAGCGGAGCAGTTAGAGCGGT 2550
QY 241 CATCAGCCCACTCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGGGACTCAGAGAC 300
DB 2551 CGTCGCCCACTCCCAACAGACACTTGGGTTTCTCTGTTGAGAGGGGGGACTCAGAGAC 2610
QY 301 AGGACTAGCTGGATTTCTTAGGCCCAACAGAAATCCCTAAGCCTAGCTGGGAAGTGACT 360
DB 2611 AGGACTAGCTGGATTTCTTAGGCCCTAGCTAAGAAATCCCTAAGCCTAGCTGGGAAGTGACT 2670
QY 361 GCATCCACTCTAAACATGGGGCTTGCNACTTAGCTACACCCGCAATCAGAGAGCTC 420
DB 2671 ACATCCACTTTAAACACGGGCTTGCNACTTAGCTACACCTGACCAATCAGAGAGCTC 2730
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCCAATCATCTATTGCCTG 480
DB 2731 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCCAATCATCTATTGCCTG 2790
QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATATAACCCAGGCATTCGAGCGGCAACGG 540
DB 2791 AGAGCACAGCGGAGGAGCAATGATCGGGATATAAACCCAAAGTCTTCGAGCGGCAACGG 2850
QY 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTTCACCTATTTCACCTCT 600
DB 2851 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTTCACCTATTTCACCTCT 2910
QY 601 ATTAATCTTCACTGAAATGAAAAAAGAAAAA 635
DB 2911 ATTAATCTTCACTGCAAAAAAAGAAAAA 2945

RESULT 16
ASK89296/c
ID ASK89296 standard; DNA; 326014 BP.
XX AC ASK89296;
XX DT 16-OCT-2002 (first entry)
XX DE Human gene for novel serine/threonine serine kinase.
XX Human; ds; gene; serine/threonine protein kinase; inflammation; cancer;
KW

```
KW arteriosclerosis; psoriasis; SNF kinase; transgenic; chromosome 3; SNP;
XX single nucleotide polymorphism.
XX Homo sapiens.
XX Location/Qualifiers
FT Key replace(1139..1140,GGT)
FT variation /*tag= ai
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(1399,G)
FT /*tag= aj
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(1871..1873,TG)
FT /*tag= ak
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(1875..1877,TA)
FT /*tag= al
FT /*standard_name= "Single nucleotide polymorphism"
FT CDS 3000..323016
FT /*tag= a
FT /*product= "Serine/threonine protein kinase"
FT exon 3000..3169
FT /*tag= b
FT FT number= 1
FT intron 3170..53737
FT /*tag= c
FT number= 1
FT variation replace(3180,G)
FT /*tag= am
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(7590,A)
FT /*tag= an
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(9315,C)
FT /*tag= ao
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(9572,A)
FT /*tag= ap
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(11004,G)
FT /*tag= aq
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(11379..11380,AAA)
FT /*tag= ar
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(12478,A)
FT /*tag= as
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(12486..12487,TTT)
FT /*tag= at
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(13904,G)
FT /*tag= au
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(25444..25446,TT)
FT /*tag= av
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(26805,A)
FT /*tag= aw
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(28010,R)
FT /*tag= ax
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(28107,Y)
FT /*tag= ay
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(28178,R)
FT /*tag= az
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(28356,T)
FT /*tag= ba
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(29535,G)
FT /*tag= bb
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(31808,T)
FT /*tag= bc
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(32807,C)
FT /*tag= bd
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(34531,G)
FT /*tag= be
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(35318,G)
FT /*tag= bf
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(35463,C)
FT /*tag= bg
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(35477,T)
FT /*tag= bh
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(37050,G)
FT /*tag= bi
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(37470,G)
FT /*tag= bj
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(37481,A)
FT /*tag= bk
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(37499,T)
FT /*tag= bl
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(40178..40179,GRA)
FT /*tag= bm
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(40787,A)
FT /*tag= bn
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(44582,G)
FT /*tag= bo
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(45615,A)
FT /*tag= bp
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(46091,T)
FT /*tag= bq
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(47844,C)
FT /*tag= br
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(48013,T)
FT /*tag= bs
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(48072,A)
FT /*tag= bt
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(50802,T)
FT /*tag= bu
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(53305,A)
FT /*tag= bv
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace(53378..53380,GG)
FT /*tag= bw
FT /*standard_name= "Single nucleotide polymorphism"
FT exon 53738..53903
FT /*tag= d
FT number= 2
FT intron 53904..83114
FT /*tag= e
FT number= 2
FT variation replace(56233,G)
FT /*tag= bx
```

Query Match	83.9%	Score	532.8	DB	6	Length	326014
Best Local Similarity	90.5%	Pred.	No.1.3e-153				
Matches	583	Conservative	0	Mismatches	52	Indels	9
Gaps	1						
Qy	1	CCCTGTATCTTTAAACCTCCTCTGTTAAAGTTTGTCTCTTCCAGAATCAAAACTCTAAACTTA	60				
Db	170822	CCCTGTATCTTTAAACCTCCTCTGTTAAAGTTTGTCTCTTCCAAAATTGAAGCTCTAAAAC	74				
Qy	61	CAAAATTGTTCTTTCAAATGGAGCACCGATGGAGTCCATGACTTAAGATCCACCGTGGACCC	120				
Db	170762	CAAAATTGTTCTTTCAAATGGAGCCCCAGACCGCAGTCCATGACTTAAGATCTTACCGCAGAC	170703				
Qy	121	CTGGACCGGCTGTCTAGCCCATGCTCCGATGTTAATGACATTGAAGGCACCCCTCCCGAG	180				
Db	170702	CTGGACCGGCTGTCTAGCCCATGCTCCGATGTTAATGACATCGAAGGCACCCCTCCAGAG	170643				
Qy	181	GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTTCAGCGGGAAGCAGTTAGACGGT	240				
Db	170642	GAAATCTCAACTGCATTAACCCCACTTACGCCCCCAATTTCAGCAGGAAGCAGTTAGACAGT	170583				
Qy	241	CATCAGCCAACTCTCCCAACAGCAGCTTGGGTTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	300				
Db	170582	CATCGGCCATCTCTCCCAACAGCAGCTTGGGTTTTTCTGTTGAGAGGGGGTACTTCAGAGAC	170523				
Qy	301	AGGACTAGCTGGAATTTCTTAGGCCAACGAAAGATCCCTAAGCCCTACTCGGAAGGTGACT	360				
Db	170522	AGGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCCTAAGCCCTACTCGGAAGGTGACT	170463				

Qy	361	GCATCCACTCTTAAACATGCGGGCTTGCAACTTAGCTCACACCCGACCACAATC-----	411
Ds	170462	GCTTCTACCTTTAAACCCCGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGGTAGGAA	170403
Qy	412	AGAGAGCTCACTAAAATGCTAAATTAGGCCAAAAATAGGAGGTAAAGAATAAGCCAATCATC	471
Ds	170402	AGAGAGCTCACTAAAATGCTAAATTAGGCCAAAAATAGGAGGTAAAGAATAAGCCAATCATC	170343
Qy	472	TATTGGCTTGAGAGCACACGCGGAGGGAACAAGGATCGGATATATAAACCCAGGCATTTCGAGC	531
Ds	170342	TATCGCTGAGAACACACAGTGGAGGGACAATGATTGGGTTATAAACCAGGCATTTCGAGC	170283
Qy	532	CGGCAACGGCAACCCCTTTGGGTCCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTTA	591
Ds	170282	CAGCAATGGTACCAATTTTTGGGTCCCCTCCCTTTGTATGGAGCTCTGTTTCACTCTTA	170223
Qy	592	TTTTCACCTATTAAATCTTGCAACTGAAAAAAAAAAAAAAAAAA	635
Ds	170222	TTTTCACCTATTAAATCTTGCAACTGAAAAAAAAAAAAAAAAAA	170179
 RESULT 17 ADQ94981/c			
ID	ADQ94981	standard; DNA; 326014 BP.	
XX	ADQ94981;		
XX	AC		
XX	AC		
DT	23-SEP-2004	(first entry)	
XX	Human kinase genomic DNA.		
DE			
XX	Kinase; diagnosis; cancer; tissue growth abnormality;		
KW	drug screening assay; cytostatic; Gene therapy; human; SNP;		
KW	single nucleotide polymorphism; gene; ds.		
XX			
XX	Homo sapiens.		
OS			
FH	Key	Location/Qualifiers	
FT	variation	replace(1138..1140,GT)	
FT		/tag= a	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(1399,G)	
FT		/tag= b	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(1871..1873,TG)	
FT		/tag= c	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(1875..1877,TA)	
FT		/tag= d	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	
FT	CDS	3000..323016	
FT		/tag= f	
FT		/product= "Human kinase protein"	
FT		/transl_except= (pos:3024..3026, aa:Arg)	
FT		/transl_except= (pos:142261..142263, aa:Glu)	
FT		/transl_except= (pos:142261..144102, aa:Glu-Ser)	
FT	exon	3000..3169	
FT		/tag= e	
FT	intron	3170..53737	
FT		/tag= g	
FT	variation	replace(3180,A)	
FT		/tag= h	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(7590,A)	
FT		/tag= i	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(9315,C)	
FT		/tag= j	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	
FT	variation	replace(9572,A)	
FT		/tag= k	
FT		/standard name= "Single nucleotide polymorphism (SNP)"	

```
FT variation replace(11004,G)
FT /tag= l
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(11378, .11380,AA)
FT /tag= m
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(12478,A)
FT /tag= n
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(12485, .12487,TC)
FT /tag= o
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(13904,G)
FT /tag= p
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(15444, .25446,TT)
FT /tag= q
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(16805,A)
FT /tag= r
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28010,R)
FT /tag= s
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28107,Y)
FT /tag= t
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28178,R)
FT /tag= u
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(28356,T)
FT /tag= v
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(29535,G)
FT /tag= w
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(31808,T)
FT /tag= x
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(32807,C)
FT /tag= y
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(33531,G)
FT /tag= z
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(35318,G)
FT /tag= aa
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(35463,G)
FT /tag= ab
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(35477,T)
FT /tag= ac
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37050,G)
FT /tag= ad
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37470,G)
FT /tag= ae
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37481,A)
FT /tag= af
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(37499,T)
FT /tag= ag
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(40177, .40179,AA)
FT /tag= ah
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(40178,A)
FT /tag= ai
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(40787,A)
FT /tag= aj
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(44582,G)
FT /tag= ak
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(45615,A)
FT /tag= al
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(46091,T)
FT /tag= am
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(47844,C)
FT /tag= an
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(48013,T)
FT /tag= ao
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(48072,A)
FT /tag= ap
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(50802,T)
FT /tag= aq
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(53305,A)
FT /tag= ar
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(53378, .53380,GG)
FT /tag= as
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT 53738, .53903
FT /tag= at
FT 53904, .83114
FT /tag= au
FT replace(56233,G)
FT /tag= av
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(56453,C)
FT /tag= aw
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(59300,T)
FT /tag= ax
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(60998,A)
FT /tag= ay
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(63013,A)
FT /tag= az
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(63293,A)
FT /tag= ba
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(64181,T)
FT /tag= bb
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(64191,A)
FT /tag= bc
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(68992,C)
FT /tag= bd
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(76246,Y)
FT /tag= be
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(7368,G)
FT /tag= bf
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(78043,A)
FT /tag= bg
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(79395,T)
FT /tag= bh
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(79474,G)
FT /tag= bi
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation
```



```
FT      /*tag= bi
XX      /standard_names "Single nucleotide polymorphism (SNP)"
PN      replace(79992,T)
FT      /*tag= bj
FT      /standard_name= "single nucleotide polymorphism (SNP)"

Query Match      83.9%; Score 532.8; DB 12; Length 326014;
Best Local Similarity 90.5%; Pred. No. 1.3e-153;
Matches 583; Conservative 0; Mismatches 52; Indels 9; Gaps 1;

QY      1      CCTGTATCTTTAACTCTCTGTTAAAGTTGTCTCTCCAGAAATCAAAACGTAAACTA 60
DB      170822 CCTGTATCTTTAACTCTCTGTTAAAGTTGTCTCTCCAAATTTGAAGCTCTAAACTA 170763

QY      61      CAAATTTGTTCTTCAATGAGACACAGATGAGTCCATGACTTAAGATCCACCGTGGAGCC 120
DB      170762 CAAATTTGTTCTTCAATGAGACACAGATGAGTCCATGACTTAAGATCTACCGAGAGCC 170703

QY      121     CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGATTAAGGACACCCCTCCCGAG 180
DB      170702 CTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGATTAAGGACACCCCTCCAGAG 170643

QY      181     GAAATCTCAACTGACCAACCCCTACTATGCCCCAATTGAGCGGAGAGAGTGTAGAGCGGT 240
DB      170642 GAAATCTCAACTGACCAACCCCTACTATGCCCCAATTGAGCGGAGAGAGTGTAGAGCGGT 170583

QY      241     CATCAGCCAACTCTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
DB      170582 CATCAGCCAACTCTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGGTACTCAGAGAC 170523

QY      301     AGACTAGCTGGATTTCTAGGCCACGAGAGATCCCTAAGCCTAGCTGGGAGGTGACT 360
DB      170522 AGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAGGTGACT 170463

QY      361     GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCAACCCGACCAATC----- 411
DB      170462 GCTTCTACCTTTAAACCCGGGCTTGCAACTTAGCTCAACCTGACCAATCAGTAGGAA 170403

QY      412     AGAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATC 471
DB      170402 AGAGAGCTCACTAAATGCTAATAGGCTAAACAGAGGTAAAGAAATAGCAATCATC 170343

QY      472     TATGCTGAGACACAGCGGGAGGACAGAGATCGGGATATAAACCCAGGCAATTCGAGC 531
DB      170342 TATGCTGAGACACAGCGGGAGGACAGATGATGGGTTATAAACCCAGGCAATTCGAGC 170283

QY      532     CGGCAACGGCAACCCCTTTGGGTCCCTCTTGTATGGGCGCTCTGTTTCACTCTA 591
DB      170282 CAGCAATGGCTACCATTTTGGGTCCCTCTTGTATGGGAGCTCTGTTTCACTCTA 170223

QY      592     TTTCACTCTATTAATCTGCAACTGAAAAAAGAAAAAAGAAAAA 635
DB      170222 TTTCACTCTATTAATCTGCAACAGCAAAAAAAGAAAAAAGAAAAA 170179
```

```
RESULT 18
AAF55630
ID      AAF55630 standard; DNA; 2781 BP.
AC      AAF55630;
XX      29-MAY-2001 (first entry)
XX      Nucleotide sequence of a human endogenous retrovirus envelope protein.
DE      Envelope protein; HERV; syncytia formation; placental development;
KW      syncytia; cancer; cell adhesion; ss.
XX      Human endogenous retrovirus.
OS      Location/Qualifiers
FH      762..2378
FT      /*tag= a
```

```
FT      /product= "envelope protein"
XX      WO200116171-A1.
PN      08-MAR-2001.
XX      01-SEP-2000; 2000WO-FR002429.
XX      01-SEP-1999; 99FR-00011141.
PR      15-SEP-1999; 99FR-00011793.
XX      (INMR ) BIO MERIEUX.
PA      (INMR ) INST NAT SANTE & RECH MEDICALE.
XX      Mallet F, Cosset F, Blond J, Lavillette D, Bouton O, Ruggieri A;
DR      WPI; 2001-226676/23.
DR      P-PSDB; AAB67652.
XX      Detecting expression of human endogenous retrovirus envelope protein in
XX      cells of a tissue or culture, from its ability to induce syncytia.
PS      Disclosure; Page 44-45; 57pp; French.
XX      The present sequence encodes a human endogenous retrovirus envelope
XX      protein. The specification describes a method for detecting expression of
XX      an envelope protein from a human endogenous retrovirus (HERV), in cells,
XX      of a tissue or culture. The method comprises detecting syncytia formation
XX      due to the fusogenic properties of the envelope protein. Envelope
XX      polypeptides and polynucleotides are used to produce therapeutic or
XX      prophylactic compositions, particularly for treatment of cancer, to
XX      correct defects in placental development for other natural formation of
XX      other types of syncytia), and to promote adhesion of cells in grafts or
XX      cellular repair processes. Expression of sequences antisense to the
XX      polynucleotide are used to prevent formation of syncytia
XX      Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 U; 0 Other;

Query Match      83.8%; Score 532.2; DB 5; Length 2781;
Best Local Similarity 91.3%; Pred. No. 2e-154;
Matches 580; Conservative 0; Mismatches 43; Indels 12; Gaps 1;

QY      1      CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGAAATCAAAACGTAAACTA 60
DB      2157 CCCTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGAAATCGAAGCTGTAAACTA 2216

QY      61      CAAATTTGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTTAAGATCCACCGTGGAGCC 120
DB      2217 -----CAAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGAGCC 2264

QY      121     CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 180
DB      2265 CTGGACCGGCTGCTAGCCCATGCTGATGTTAATGACATCAAGAGGACCCCTCTGAG 2324

QY      181     GAAATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB      2325 GAAATCTCAGTGCACAACTCTACTACGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 2384

QY      241     CATCAGCCAACTCTCCCAACAGACATTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
DB      2385 CGTGGCCCAACTCTCCCAACAGACATTTAGGTTTCTGTTGAGATGGGGGACTGAGAGAC 2444

QY      301     AGGACTAGCTGGATTTCTAGCCCAACGAAGAAATCCCTAAGCTAGCTGGGAAGGTGACT 360
DB      2445 AGGACTAGCTGGATTTCTAGCTGACTAAGAAATCCCTAAGCTAGCTGGGAAGGTGACT 2504

QY      361     GCATCCACCTCTAAACATGGGGCTTGGCAACTTAGCTCAACCCGACCAATCAGAGAGCTC 420
DB      2505 ACATCCACCTTTAAACACGGGGCTTGGCAACTTAGCTCAACCTGACCAATCAGAGAGCTC 2564

QY      421     ACTAAAAATGCTAATTAGGCAAAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCTCG 480
DB      2565 ACTAAAAATGCTAATTAGGCAAAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCTCG 2624
```


QY 481 AGGACACAGGAGGAGGACGAGATCGGATATATACCCAGGCAATTCGAGCCGGCAACGG 540
 Db 2625 AGAGCACAGGAGGAGGACGAGATCGGATATATACCCAGGCAATTCGAGCCGGCAACGG 2684
 QY 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACTCT 600
 Db 2685 CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACTCT 2744
 QY 601 ATTAATCTTGCACCTGAAAAAATAAAAAA 635
 Db 2745 ATTAATCTTGCACCTGAAAAAATAAAAAA 2779

RESULT 19

AA25660
 ID AAX25660 standard; cDNA to mRNA; 1136 BP.

AC AAX25660;

DT 21-MAY-1999 (first entry)

DE Human endogenous retrovirus W clone cl.C4C5.

XX Clone; human endogenous retrovirus; genome; autoimmune disease;
 KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
 KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.

XX Human endogenous retrovirus.

OS WO9902696-A1.

PN 21-JAN-1999.

PD 06-JUL-1998; 98WO-FR001442.

PF 07-JUL-1997; 97FR-00008815.

PR (INMR) BIO MERIEUX.

XX Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;

XX WPI; 1999-120897/10.

XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
 PT disease, and abnormal or failed pregnancy.

XX Claim 1; Page 59-60; 106pp; French.

XX This sequence represents clone cl.C4C5 of the human endogenous retrovirus
 CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
 CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
 CC dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility

XX Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;

Query Match 82.8%; Score 525.8; DB 2; Length 1136;
 Best Local Similarity 90.7%; Pred. No. 1.3e-152;
 Matches 576; Conservative 0; Mismatches 47; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTTGTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAACTA 60
 Db 510 CCCTGTATCTTTAACTCTTGTGTTAAAGTTTCTCTTCCAGAAATCGAGCTGTAAACTA 569
 QY 61 CAATTTGTTCTTAAATGGAGCACGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
 Db 570 -----CAATGGAGCCCAAGATGCAGTCCAAGACTAAGATCTACCGCAGACCC 617

QY 121 CTGGACCGGCTCTGTAGCCCATGCTCCGATGTATATGACATTTGAAGGCAACCCCTCCCGAG 180
 Db 618 CTGGACCGGCTCTGTAGCCCATGCTCCGATGTATATGACATTTGAAGGCAACCCCTCCCGAG 677
 QY 181 GAAATCTCACTGCAACACCCCTACTATGCCCCCAATTCAGCGGGAAGAGTGTAGAGCGGT 240
 Db 678 GAAATCTCACTGCAACACCCCTACTATGCCCCCAATTCAGCGGGAAGAGTGTAGAGCGGT 737
 QY 241 CATCAGCCAACTCCCCAAACAGCAGCTTTGGGTTTCTGTTTGTAGAGGGGGGACTGTAGAGAC 300
 Db 738 CGTCGGCCAACTCCCCAAACAGCAGCTTTAGGTTTCTGTTTGTAGATGGGAGTGTAGAGAC 797
 QY 301 AGGACTAGCTGGATTTCTTAGGCCAAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 798 AGGACTAGCTGGATTTCTTAGGCCAAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 857
 QY 361 GCATCCACCTCTAATACATGGGCTTGCACCTTAGCTCACACCCGACCAATTCAGAGAGCTC 420
 Db 858 ACATCCACCTTTAAACACCGGGCTTGCACCTTAGCTCACACCTGACCAATTCAGAGAGCTC 917
 QY 421 ACTAAATGCTAATTAGGCAAAATAAGAGGTAAGAAATAGCAATTCATCTATTGCTCTG 480
 Db 918 ACTAAATGCTAATTAGGCAAAATAAGAGGTAAGAAATAGCAATTCATCTATTGCTCTG 977
 QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGATATATAACCCAGGCAATTCGAGCCGGCAACGG 540
 Db 978 AGAGCACAGCGGAGGAGCAAGGATCGGATATATAACCCAGGCAATTCGAGCCGGCAACGG 1037
 QY 541 CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACTCT 600
 Db 1038 CAACCCCTTTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTTCACTCT 1097
 QY 601 ATTAATCTTGCACCTGAAAAAATAAAAAA 635
 Db 1098 ATTAATCTTGCACCTGAAAAAATAAAAAA 1132

RESULT 20

AA59210
 ID AA59210 standard; DNA; 1136 BP.

AC AA59210;

DT 07-NOV-2000 (first entry)

DE 3' pol gene and 3' non coding sequences of HERV-W from human genome.

XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
 KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX Homo sapiens.

XX WO200043521-A2.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-FR000144.

XX 21-JAN-1999; 99FR-00000888.

XX (INMR) BIO MERIEUX.

XX Paranhos-Baccala G, Mallet F, Voisset C;

XX WPI; 2000-499229/44.

XX New nucleic acid from human endogenous retrovirus, useful e.g. for
 PT diagnosis of autoimmune disease and complications of pregnancy, contains
 PT at least part of the gag gene.

XX Disclosure; Page 46; 53pp; French.

XX The present sequence represents an endogenous retroviral nucleic acid

CC	fragment, which is associated with an autoimmune disease, and is	
CC	integrated into the human genome. The fragment is originally derived from	
CC	a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W	
CC	retrovirus is associated with autoimmune disease, failure of pregnancy or	
CC	disorders of pregnancy. The nucleic acid fragment, or proteins derived	
CC	from it, are useful for diagnosis of autoimmune disease (specifically	
CC	multiple sclerosis) and for monitoring pregnancy. The nucleic acid	
CC	fragments may also be used for in situ labelling of isolated chromosomes,	
CC	while the transcription product can be used to study or monitor T cell	
CC	proliferation in vitro	
XX		
XX	Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;	
SQ		
	Query Match 82.8%; Score 525.8; DB 3; Length 1136;	
	Best Local Similarity 90.7%; Pred. NO. 1.3e-152;	
	Matches 576; Conservative 0; Mismatches 47; Indels 12; Gaps 1;	
QY	1 CCTGTATCTTTAACTCTCTTGTAGTTTGTCTCTCCAGAACTCAAAACTGTAAACTA 60	
Db		
Db	510 CCTGTATCTTTAACTCTCTTGTAGTTTGTCTCTCCAGAACTCAAAACTGTAAACTA 569	
QY	61 CAAATTTGTTCTTCAAAATGGAGCACCAGATGGAGTGCATGAAGATCCACCCTGGACCC 120	
Db		
Db	570 -----CAAATGGAGCCCAAGATGCATGCAAGACTTAAGATCTACCGCAGACCC 617	
QY	121 CTGGACGGCTCTAGCCATGCTCCGATGTTAATGACATGAAGGACCCCTCCCGAG 180	
Db		
Db	618 CTGGACGGCTCTAGCCATGCTCCGATGTTAATGACATCAAAAGGCCCTCTCTGAG 677	
QY	181 GAATCTCACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240	
Db		
Db	678 GAAATCTCAGCTGCACAACTCTACTACGCCCCAAATTCAGCAGGAAGCAGTTAGACGGT 737	
QY	241 CATCAGCCAACTCCCAACAGACATTTGGGTTTCTCTGTGAGAGGGGAGCTGAGAGAC 300	
Db		
Db	738 CGTCGGCCAACTCCCAACAGACATTTAGTTTCTCTGTGAGATGGGGACTGAGAGAC 797	
QY	301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCTAGCTGGAGGTGACT 360	
Db		
Db	798 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGCC 857	
QY	361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATTCAGAGAGCTC 420	
Db		
Db	858 ACATCCACCTTTTAAACACGGGGCTTGCACTTAGTTTCAACCTGACCAATTCAGAGAGCTC 917	
QY	421 ACTAAATGCTAATTAGGCAAAATAGGAGTTAAGAAATAGCCAAATCATCTATTGCGCTG 480	
Db		
Db	918 ACTAAATGCTAATTAGGCAAAACAGACAGGAGTTAAGAAATAGCCAAATCATCTATTGCAATG 977	
QY	481 AGACACAGCGGGAGGAGACAGGATCGGGATATAAACCCAGGCAATTCGAGCGCGCAACGG 540	
Db		
Db	978 AGACACAGCAGGAGGAGCAATGATCGGGATATAAACCCAAAGTCCTCGAGCGCGCAACGG 1037	
QY	541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600	
Db		
Db	1038 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCATGCTATTTCACCTCT 1097	
QY	601 ATTAATTTGCAACTGAAAAAATAAAAAA 635	
Db		
Db	1098 ATTAATTTGTCAGCTCGAAAAAATAAAAAA 1132	
	RESULT 21	
	AAAX25661	
ID	AAAX25661 standard; cDNA to mRNA; 2782 BP.	
XX		
AC	AAAX25661;	
XX		
DT	21-MAY-1999 (first entry)	
XX		
DE	Human endogenous retrovirus W clone cl.PH74.	
XX		
XX	Clone; human endogenous retrovirus; genome; autoimmune disease;	
KW		

KW	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;	
KW	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.	
XX		
OS	Human endogenous retrovirus.	
XX	WO9902696-A1.	
PN	21-JAN-1999.	
PD	06-JUL-1998; 98WO-FR001442.	
XX	07-JUL-1997; 97FR-00008815.	
XX	(INMR) BIO MERIEUX.	
PA	Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;	
PI	WPI; 1999-120897/10.	
XX		
DR	New nucleic acid sequences from human endogenous retrovirus-W - expressed	
XX	exclusively in placenta and useful in diagnosis and therapy of autoimmune	
PT	disease, and abnormal or failed pregnancy.	
PT	Claim 1; Page 60-63; 106pp; French.	
XX		
CC	This sequence represents clone cl.PH74 of the human endogenous retrovirus	
CC	(HERV) W genome. The nucleic acids, their fragments or peptides encoded	
CC	by them are markers of autoimmune disease (e.g. multiple sclerosis,	
CC	rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-	
CC	dependent diabetes and related pathologies) and of abnormal or	
CC	unsuccessful pregnancy and can be used as chromosomal markers for	
CC	susceptibility to these conditions, or proximity markers of genes	
CC	associated with this susceptibility	
XX		
SQ	Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;	
	Query Match 82.3%; Score 522.6; DB 2; Length 2782;	
	Best Local Similarity 90.4%; Pred. NO. 2e-151;	
	Matches 574; Conservative 0; Mismatches 49; Indels 12; Gaps 1;	
QY	1 CCTGTATCTTTAACTCTCTTGTAGTTTGTCTCTCCAGAACTCAAAACTGTAAACTA 60	
Db		
Db	2158 CCTGTATCTTTAGCTCTCTTGTAACTTTGTCTCTCCAGAACTCAAAACTGTAAACTA 2217	
QY	61 CAAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC 120	
Db		
Db	2218 -----CAAATGGAGCCCAAGATGCAGTCCCAAGACTAGACTACCGCAGACCC 2265	
QY	121 CTGGACCGCTCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCCGAG 180	
Db		
Db	2266 CTGGACCGCTCTGTAGCCCATGCTCCGATGTTAATGACATCAAAAGGACCCCTCCCTGAG 2325	
QY	181 GAAATCTCACTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240	
Db		
Db	2326 GAAATCTCAGCTGCACACCCCTCTACTCGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 2385	
QY	241 CATCAGCCAACTCCCAACAGACATTTGGGTTTCTCTGTGAGAGGGGAGCTGAGAGAC 300	
Db		
Db	2386 GGTCCGCCAACCTCCCAACAGACATTTAGTTTCTCTGTGAGATGGGGACTGAGAGAC 2445	
QY	301 AGGACTAGCTGGATTTCTTAGCCCAACGAAGATCCCTAAGCTAGCTGGGAAGGTGACT 360	
Db		
Db	2446 AGGACTAGCTGGATTTCTTAGCTGACTAAGAAATCTTAAAGCTAGTGGGAAGGTGACC 2505	
QY	361 GCATCCACCTCTAAACATGGGCTTGGCACTTAGCTCACCCCGACCAATTCAGAGAGCTC 420	
Db		
Db	2506 ACATCCACCTTTAAACACGGGGCTTGGCACTTAGCTTCACCTGACCAATTCAGAGAGCTC 2565	
QY	421 ACTAAAAATGCTAATTAGGCAAAATAGGAGTTAAGAAATAGCCCAATCATCTATTGCGCTG 480	
Db		
Db	2566 ACTAAAAATGCTAATTAGGCAAAAGAGGAGTTAAGAAATAGCCCAATCATCTATTGCGCTG 2625	
QY	481 AGACACAGCGGGAGGAGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCGCGCAACGG 540	

Db 2626 AGAGCACAGCAGGAGGACATGATCGGATATAACCCAGTTTTCGAGCCGGCAACGG 2685
QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACTCT 600
Db 2686 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACTCT 2745
QY 601 ATTAAATCTTGCACCTCAAAAAAAGAAAAA 635
Db 2746 ATTAAATCTTGCACCTCAAAAAAAGAAAAA 2780

RESULT 22

AAH59211 ID AAH59211 standard; DNA; 2782 BP.

XX AC

XX AC

XX DT 07-NOV-2000 (first entry)

XX XX

XX DE 5' non coding, 3' pol gene, and 3' non coding sequences of HERV-W.

XX KW Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;

XX KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.

XX XX

XX OS Homo sapiens.

XX XX

XX PN WO200043521-A2.

XX XX

XX PD 27-JUL-2000.

XX XX

XX PF 21-JAN-2000; 2000WO-FR000144.

XX XX

XX PR 21-JAN-1999; 99FR-00000888.

XX XX

XX PA (INMR) BIO MERIEUX.

XX XX

XX PI Paranhos-Baccala G, Mallet F, Voisset C;

XX XX

XX DR WPI; 2000-499229/44.

XX XX

XX PT New nucleic acid from human endogenous retrovirus, useful e.g. for
XX PT diagnosis of autoimmune disease and complications of pregnancy, contains
XX PT at least part of the gag gene.

XX XX

XX PS Disclosure; Page 46-47; 53pp; French.

XX XX

XX CC The present sequence represents an endogenous retroviral nucleic acid
XX CC fragment, which is associated with an autoimmune disease, and is
XX CC integrated into the human genome. The fragment is originally derived from
XX CC a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W or
XX CC retrovirus is associated with autoimmune disease, failure of pregnancy or
XX CC disorders of pregnancy. The nucleic acid fragment, or proteins derived
XX CC from it, are useful for diagnosis of autoimmune disease (specifically
XX CC multiple sclerosis) and for monitoring pregnancy. The nucleic acid
XX CC fragments may also be used for in situ labelling of isolated chromosomes,
XX CC while the transcription product can be used to study or monitor T cell
XX CC proliferation in vitro

XX SQ Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 82.3%; Score 522.6; DB 3; Length 2782;
Best Local Similarity 90.4%; Pred. No. 2e-151;
Matches 574; Conservative 0; Mismatches 49; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAACTA 60

Db 2158 CCCTGTATCTTTGACCTCTCTTTAACTTTCTCTTCCAGAAATCGAGCTGTAAACTA 2217

QY 61 CAAATTTGTTTCAAAATGGAGCACCGATGGAGTCCATGATTAAGATCCACCGTGGACCC 120

Db 2218 -----CAATGGAGCCCAAGATGATGAGTCCAAGACTAAGATCTACCGGAGACCC 2265

QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGATTAAGGACACCCCTCCCGAG 180
Db 2266 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGATTAAGGACACCCCTCCCGAG 2325
QY 181 GAAATCTCACTGACCAACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTTAGACGGT 240
Db 2326 GAAATCTCACTGACCAACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTTAGACGGT 2385
QY 241 CATCAGCCAAACCTCCCCCAACAGCAGCTTGGGTTTTCCTGTTGAGAGGGGGAGCTGAGAGAC 300
Db 2386 GGTGGCCAACTCCCCCAACAGCAGCTTGGGTTTTCCTGTTGAGAGGGGGAGCTGAGAGAC 2445
QY 301 AGGACTAGCTGATTTCTTAGGCCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 2446 AGGACTAGCTGATTTCTTAGGCCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 2505
QY 361 GCATCCACCTCTAAGACATGGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 420
Db 2506 ACATCCACCTTTAAACACAGGGGCTTGCACCTTAGCTCAGCCGACCAATCAGAGAGCTC 2565
QY 421 ACTAAATCTTAATTAGGCAAAAAATAGGAGGTAAAGAAATAGCCAATCATTTATTCCTG 480
Db 2566 ACTAAATCTTAATTAGGCAAAAAATAGGAGGTAAAGAAATAGCCAATCATTTATTCCTG 2625
QY 481 AGAGCACAGCGGAGGACAAAGATCGGATATAAACCCAGGATTCGAGCCCGGCAACGG 540
Db 2626 AGAGCACAGCGGAGGACAAATGATCGGATATAAACCCAGGATTCGAGCCCGGCAACGG 2685
QY 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACTCT 600
Db 2686 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACTCT 2745
QY 601 ATTAAATCTTGCACCTCAAAAAAAGAAAAA 635
Db 2746 ATTAAATCTTGCACCTCAAAAAAAGAAAAA 2780

RESULT 23

AAH20069

ID AAH20069 standard; DNA; 2782 BP.

XX AC

XX AC

XX DT 08-AUG-2001 (first entry)

XX XX

XX DE HERV-W envelope protein G encoding nucleic acid.

XX XX

XX KW Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
XX KW envelope protein; multiple sclerosis-related superantigen; vaccine;
XX KW surface antigen; transmembrane; multiple sclerosis; neuroprotective;
XX KW antisense-therapy; autoimmune disorder; ds.

XX OS

XX XX

XX FH Human endogenous retrovirus.

XX FT

XX FT 5'UTR

XX FT

XX FT CDS

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

3' UTR	/transl_except= (pos:2170..2172,aa:Asn)	2386	GGTCGGCAACCTCCCAACACAGACACTTAGGTTTCTCTGTTAGATGGGGGACGTGAGAGAC	2444
FT	2380..2782	301	AGGACTAGCTGGATTTCTTAGGCAACCAAGAAATCCTTAAGCCTAGCTGGGAAAGGTGACT	360
FT	/*tag= C	2446	AGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCTTAAGCCTAGCTGGGAAAGGTGACC	2505
XX		361	GCATCCACCTCTAAACATGCGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC	420
XX		2506	ACATCCACCTTTAAACACGCGGCTTGCACTTAGCTCACACCCGACCAATCAGAGAGCTC	2565
XX		421	ACTAAATGCTAATTAGGCAAAATAGAGTAAGAAATAGCAATCATCTATTTCCTTG	480
XX		2566	ACTAAATGCTAATTAGGCAAAATAGAGTAAGAAATAGCAATCATCTATTTCCTTG	2625
XX		481	AGAGCACAGCGGAGGACCAAGGATCGGATATAAACCAGGATTCGAGCCGCGCAACGG	540
XX		2626	AGAGCACAGCAGGAGGACCAATGATCGGATATAAACCAGGATTCGAGCCGCGCAACGG	2685
XX		541	CAACCCCTTTGGTCCCTCCCTTTCTGATGGGCGCTCTGTTTTCACCTCTATTTCACCTCT	600
XX		2686	CAACCCCTTTGGTCCCTCCCTTTCTGATGGGCGCTCTGTTTTCACCTCTATTTCACCTCT	2745
XX		601	ATTAATCTTGCACCTGAAAAAATAAAAAAATAAAAAA 635	
XX		2746	ATTAATCTTGCACCTGAAAAAATAAAAAAATAAAAAA 2780	
XX		RESULT 24		
XX		AAD24195		
ID		AAD24195 standard; cDNA; 2930 BP.		
XX		AAD24195;		
XX		07-MAY-2002 (first entry)		
XX		Human syncytin cDNA.		
XX		Human; syncytin; preeclampsia; gestational trophoblast disorder;		
XX		choriocarcinoma; hydatiform mole; placental site tumour; abortion;		
XX		envelope gene; human endogenous defective retrovirus; HERV-W; ss.		
XX		Homo sapiens.		
XX		Key Location/Qualifiers		
XX		CDS 930..2546		
XX		/*tag= a		
XX		/product= "Syncytin"		
XX		WO200204678-A2.		
XX		17-JAN-2002.		
XX		09-JUL-2001; 2001WO-US021719.		
XX		07-JUL-2000; 2000US-0216657P.		
XX		(GEM) GENETICS INST INC.		
XX		Keith JC, Mccoy JM, Mi S;		
XX		WPI; 2002-171727/22.		
XX		P-PSDB; AAE14540.		
XX		Identifying a compound for treating a subject with or at risk of		
XX		developing preeclampsia, comprises determining whether the expression or		
XX		activity of syncytin in the cell is modulated in the presence of a test		
XX		compound.		
XX		Disclosure; Page 39-42; 43pp; English.		
XX		The invention relates to identifying compounds which are modulators of		
XX		syncytin expression. The syncytin modulators are useful in diagnosis and		
XX		treatment of preeclampsia and gestational trophoblast disorders (e.g.		
XX				

CC choriocarcinoma, hydatiform mole, placental site tumour and missed/
CC incomplete abortion). Syncytin is a human gene derived from the envelope
CC gene of human endogenous defective retrovirus, HERV-W. The present
CC invention is based partly on the discovery that syncytin expression is
CC dramatically reduced in preclampsia, and is also mis-localised to the
CC apical syncytiotrophoblast membrane. The present sequence is human
CC syncytin cDNA

XX SQ Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;

Query Match 81.5%; Score 517.4; DB 6; Length 2930;
Best Local Similarity 91.4%; Pred. No. 8.3e-150;
Matches 564; Conservative 0; Mismatches 41; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTCTTGTAGTTTGTCTTCCAGAAATCAAACTGTAAACTA 60
DB |||||
QY 61 CAAATTTCTTCAATGGAGCACCAGATGGAGTCCATGATCAAGATCCACCTGACCC 120
DB |||||
QY 121 CTGACCGGCTGTAGCCCATGCTCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 180
DB |||||
QY 2433 CTGGACCGGCTGTAGCCCATGCTGTATGTTTAATGACATCAAGCACCCTCTCTGAG 2492
DB |||||
QY 181 GAATCTCACTCAGCAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGT 240
DB |||||
QY 2493 GAAATCTCAGCTGACACACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGT 2552
DB |||||
QY 241 CATCAGCAACCTCCCAACAGCAGCATTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB |||||
QY 2553 CGTCGGCAACCTCCCAACAGCAGCATTTGGGTTTTCTGTTGAGATGGGAGCTGAGAGAC 2612
DB |||||
QY 301 AGGACTAGCTGGATTTCTAGGCAACAGAAATCCCTAAGCCCTAGCTGGAAAGTGACT 360
DB |||||
QY 2613 AGGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCTAAGCCCTAGCTGGAAAGTGACC 2672
DB |||||
QY 361 GCATCCACCTTAAACATGGGCTTCACTAGCTCAGCCGACCAATCAGAGAGCTC 420
DB |||||
QY 2673 ACATCCACCTTAAACAGGGGCTTCACTAGCTCAGCCGACCAATCAGAGAGCTC 2732
DB |||||
QY 421 ACTAAATGCTAATAGGCAAAATAGAGTAAAGAAATAGCAATCATCTATTGCGCTG 480
DB |||||
QY 2733 ACTAAATGCTAATAGGCAAAACAGAGGTAAGAAATAGCAATCATCTATTGCGCTG 2792
DB |||||
QY 481 AGAGCAGCGGGAGGAGCAAGATCGGGATATAAACCCAGGATTCGAGCCGGCAACGG 540
DB |||||
QY 2793 AGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAGGATTCGAGCCGGCAACGG 2852
DB |||||
QY 541 CAACCCCTTTGGTCCCTCCCTCTTGTATGGGCGCTCTTTTCACTCTATTTCACCTCT 600
DB |||||
QY 2853 CAACCCCTTTGGTCCCTCCCTCTTGTATGGGAGGCTCTGTTTTCATCTATTTCACCTCT 2912
DB |||||
QY 601 ATTAATCTTGCAACTG 617
DB |||||
QY 2913 ATTAATCTTGCAACTG 2929
DB |||||

RESULT 25
ABL61744
ID ABL61744 standard; DNA; 56093 BP.
XX AC ABL61744;
XX DT 15-MAY-2002 (first entry)
XX DE Colon adenocarcinoma related gene sequence SEQ ID NO:81.
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.
OS WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
XX 18-SEP-2000; 2000US-0231133P.
XX 18-SEP-2000; 2000US-0233617P.
XX 20-SEP-2000; 2000US-0234009P.
XX 20-SEP-2000; 2000US-0234034P.
XX 22-SEP-2000; 2000US-0234052P.
XX 22-SEP-2000; 2000US-0234509P.
XX 22-SEP-2000; 2000US-0234567P.
XX 25-SEP-2000; 2000US-0234923P.
XX 25-SEP-2000; 2000US-0234924P.
XX 25-SEP-2000; 2000US-0235077P.
XX 25-SEP-2000; 2000US-0235082P.
XX 25-SEP-2000; 2000US-0235134P.
XX 25-SEP-2000; 2000US-0235280P.
XX 26-SEP-2000; 2000US-0235637P.
XX 26-SEP-2000; 2000US-0235638P.
XX 27-SEP-2000; 2000US-0235711P.
XX 27-SEP-2000; 2000US-0235720P.
XX 27-SEP-2000; 2000US-0235840P.
XX 27-SEP-2000; 2000US-0235863P.
XX 28-SEP-2000; 2000US-0236028P.
XX 28-SEP-2000; 2000US-0236032P.
XX 28-SEP-2000; 2000US-0236033P.
XX 28-SEP-2000; 2000US-0236034P.
XX 28-SEP-2000; 2000US-0236109P.
XX 28-SEP-2000; 2000US-0236112P.
XX 29-SEP-2000; 2000US-0236842P.
XX 29-SEP-2000; 2000US-0236891P.
XX 02-OCT-2000; 2000US-0237172P.
XX 02-OCT-2000; 2000US-0237173P.
XX 02-OCT-2000; 2000US-0237278P.
XX 02-OCT-2000; 2000US-0237294P.
XX 02-OCT-2000; 2000US-0237295P.
XX 02-OCT-2000; 2000US-0237316P.
XX 03-OCT-2000; 2000US-0237425P.
XX 03-OCT-2000; 2000US-0237598P.
XX 03-OCT-2000; 2000US-0237604P.
XX 03-OCT-2000; 2000US-0237606P.
XX 03-OCT-2000; 2000US-0237608P.
XX 01-NOV-2000; 2000US-0244867P.
XX 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX Claim 1; SEQ ID NO 81; 44pp; English.
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, adenocarcinoma, papillary carcinoma and Wilms
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
SQ Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;

Query Match 81.3%; Score 516.2; DB 6; Length 56093;
Best Local Similarity 91.1%; Pred. No. 8.3e-149;
Matches 564; Conservative 0; Mismatches 43; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCCTCTTAACTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 37274 CCTGTATCTTTAACTCCTCTTAACTTGTCTTCCAGAAATCGAAGCTGTAAACTA 37333
QY 61 CAAATTTGTTCTTCAAAATGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB 37334 -----CAATGAGCCCAAGATGCGATCCAGACTAAGATCTACCGCAGACCC 37381
QY 121 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGAGGACCCCTCCCGAG 180
DB 37382 CTGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAAAAGGACCCCTCCGAG 37441
QY 181 GAAATCTCAACTGCACAACTTACTATGCTCCCAATTCAGCGGAGCAATTAAGAGCGGT 240
DB 37442 GAAATCTCACTGCTCAAACTTACTATGCTCCCAATTCAGCGGAGCAATTAAGAGCGGT 37501
QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTCTGTGAGAGGGGGGACTGAGAGAC 300
DB 37502 CGTCGGCAACCTCCCAACAGCACTTAGGTTTCTGTGAGATGGGGGACTGAGAGAC 37561
QY 301 AGGACTAGCTGGATTTCTTAGGCCAAGAGATCCCTAAGCTAGCTGGAGGTGACT 360
DB 37562 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCTAGCTGGAGGTGACC 37621
QY 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCGGACCAATTCAGAGAGCTC 420
DB 37622 ACATCCACCTTTAAACACGGGGCTTGCACTTAGCTCACACCTGACCAATTCAGAGAGCTC 37681
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCATCACTATTGCTCG 480
DB 37682 ACTAAATGCTAATTAGGCAAAAGACAGAGGTAAAGAAATAGCCATCACTATTGCTCG 37741
QY 481 AGACACAGCGGGGAGGACAGGATCGGATATAAACCAGGCATTCGAGCCGGCAACGG 540
DB 37742 AGACACAGCGAGGAGGACAGGATCGGATATAAACCAGGATTCGAGCCGGCAACGG 37801
QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCATCTCT 600
DB 37802 CAACCCCTTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCATCTCT 37861
QY 601 ATTAATCTTGCACACTGAA 619
DB 37862 ATTAATCTTGCACACTGCA 37880

RESULT 26
AAAX25665
ID AAAX25665 standard; cDNA to mRNA; 7582 BP.
XX
AC AAAX25665;
XX
DT 21-MAY-1999 (first entry)
DE Complete human endogenous retrovirus W genome.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
XX multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW

KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX Human endogenous retrovirus.
OS
XX WO9902696-A1.
PN
XX 21-JAN-1999.
PD
XX 06-JUL-1998; 98WO-FR001442.
PF
XX 07-JUL-1997; 97FR-00008815.
PR
XX (INMR) BIO MERIEUX.
PA
XX Beesme F, Blond J, Bouton O, Mandrand B, Mallet P;
FI
XX WPI; 1999-120897/10.
DR
XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
XX exclusively in placenta and useful in diagnosis and therapy of autoimmune
XX disease, and abnormal or failed pregnancy.
PT
PT
PT
PT
PS
XX Claim 1; Page 71-74; 106pp; French.
XX
XX This sequence represents the complete sequence of the human endogenous
XX retrovirus (HERV) W genome. The nucleic acids, their fragments or
XX peptides encoded by them are markers of autoimmune disease (e.g. multiple
XX sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
XX insulin-dependent diabetes and related pathologies) and of abnormal or
XX unsuccessful pregnancy and can be used as chromosomal markers for
XX susceptibility to these conditions, or proximity markers of genes
XX associated with this susceptibility
SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;

Query Match 80.5%; Score 511.4; DB 2; Length 7582;
Best Local Similarity 89.5%; Pred. No. 9.8e-148;
Matches 552; Conservative 11; Mismatches 42; Indels 12; Gaps 1;
QY 1 CCTGTATCTTTAACTCCTCTTAACTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 6976 CCTGTATCTTTAACTCCTCTTAACTTGTCTTCCAGAAATCGAAGCTGTAAACTA 7035
QY 61 CAAATTTGTTCTTCAAAATGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB 7036 -----CAATGAGCCCAAGATGCGATCCAGACTAAGATCTACCGCAGACCC 7083
QY 121 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGAGGACCCCTCCCGAG 180
DB 7084 CTGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTAAGAGGACCCCTCCCGAG 7143
QY 181 GAAATCTCAACTGCACAACTTACTATGCCCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
DB 7144 GAAATCTCAGTGCACAACTTACTATGCCCCCAATTCAGCGGAGCAGTTAGAGCGGT 7203
QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
DB 7204 SGTGCGCAACCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGGACTGAGAGAC 7263
QY 301 AGGACTAGCTGGATTTCTTAGCCCAAGCAATCCCTAAGCTAGCTGGGAAGGTGACT 360
DB 7264 AGGACTAGCTGGATTTCTTAGCCCAAGCAATCCCTAAGCTAGCTGGGAAGGTGACT 7323
QY 361 GCATCCACCTCTAAACATGGGGCTTTGCAACTTTAGCTCTACACCGGACCAATTCAGAGAGCTC 420
DB 7324 ACATCCACCTTTAAACACGGGGCTTGCAACTTTAGTCTACACCTGACCAATTCAGAGAGCTC 7383
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCATCACTATTGCTCG 480
DB 7384 ACTAAATGCTAATTAGGCAAAAGACAGAGGTAAAGAAATAGCCATCACTATTGCTCG 7443
QY 481 AGAGCACAGCGGAGGACAGGATCGGATATAAACCAGGCATTCAGCGCGGCAACGG 540

Db 7444 AGAGCACAGGAGGACAAATGATCGGATATATAACCAAGTTCGAGCGGCAACGG 7503
Qy 541 CAACCCCTTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 7504 CAACCCCTTTGGTCCCTCCCTTTGATGGGAGCTCTGTTTCACTCTATTTCACCTCT 7563
Qy 601 ATTAATCTTGCACACTG 617
Db 7564 ATTAATCTTGCACACTG 7580

RESULT 27
AAA59215
ID AAA59215 standard; DNA; 7582 BP.
XX AC AAA59215;
XX 07-NOV-2000 (first entry)
XX Human endogenous retrovirus W (HERV-W) sequence.
XX Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
KW gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX Human endogenous retrovirus.
XX Key Location/Qualifiers
FH LTR 1..120
FT /*tag= a
FT /note= "R of 5' LTR"
FT 121..575
FT /*tag= b
FT /note= "U5 of 5' LTR"
FT primer_bind 579..596
FT /*tag= c
FT CDS 5581..7194
FT /*tag= d
FT /note= "ORF1 env538"
FT 7039..7194
FT /*tag= e
FT CDS 7112..7255
FT /note= "ORF2 52 AA"
FT /*tag= f
FT /note= "ORF3 48 AA"
FT misc_feature 7244..7254
FT /*tag= g
FT /note= "polypurine tract"
FT LTR 7256..7582
FT /*tag= h
FT polyA_signal /note= "U3-R of 3' LTR"
FT 7563..7569
FT /*tag= i

XX WO200043521-A2.
XX 27-JUL-2000.
XX 21-JAN-2000; 2000WO-FR000144.
XX 21-JAN-1999; 99FR-00000888.
XX (INNR) BIO MERIEUX.
XX Paranhos-Baccala G, Mallet F, Voisset C;
XX WPI; 2000-499229/44.
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
PT diagnosis of autoimmune disease and complications of pregnancy, contains
PT at least part of the gag gene.
XX Disclosure; Page 49-52; 53pp; French.
XX

CC The present sequence represents an endogenous retrovirus, which is
CC associated with an autoimmune disease, and is integrated into the human
CC genome. The retrovirus is human endogenous retrovirus W (HERV-W). The
CC HERV-W retrovirus is associated with autoimmune disease, failure of
CC pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or
CC proteins derived from it, are useful for diagnosis of autoimmune disease
CC (specifically multiple sclerosis) and for monitoring pregnancy. The
CC nucleic acid fragments may also be used for in situ labelling of isolated
CC chromosomes, while the transcription product can be used to study or
CC monitor T cell proliferation in vitro
XX
SQ Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;

Query Match 80.5%; Score 511.4; DB 3; Length 7582;
Best Local Similarity 89.5%; Pred. NO. 9.8e-148;
Matches 552; Conservative 11; Mismatches 42; Indels 12; Gaps 1;
Qy 1 CCTGTATCTTTAACTCCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
Db 6976 CCTGTATCTTTAACTCCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 7035
Qy 61 CAAATTTGTTCTTCAATGGAGCAGATGAGTCCATGACATGAAGATCCACCCGTGACCC 120
Db 7036 -----CAAATGGAGCCCAAGATGAGTCCATGACATGAAGATCCACCCGTGACCC 7083
Qy 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGGCACCCCTCCCGAG 180
Db 7084 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGGCACCCCTCCCGAG 7143
Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTATAGCGGT 240
Db 7144 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGAGAGCAGTATAGCGGT 7203
Qy 241 CATCAGCCCACTCCCAACAGCAGCTTGGGTTTCTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 7204 SGTGCGCAACCTCCCAACAGCAGCTTGGGTTTCTCTGTTGAGAGGGGAGCTGAGAGAC 7263
Qy 301 AGGACTAGCTGGATTTCTAGGCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 7264 AGGACTAGCTGGATTTCTAGGCCAACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 7323
Qy 361 GCATCCACCTCTAAACATGGGGCTTGCATTTAGCTCAACCCGACCAATCAGAGAGCTC 420
Db 7324 ACATCCACCTTTAAACACGGGGCTTGCATTTAGCTCAACCCGACCAATCAGAGAGCTC 7383
Qy 421 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
Db 7384 ACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 7443
Qy 481 AGAGCACAGCGGAGGACAAAGGATCGGATATAAACCCAGGATTCGAGCCGCAACGG 540
Db 7444 AGAGCACAGCGGAGGACAAAGGATCGGATATAAACCCAGGATTCGAGCCGCAACGG 7503
Qy 541 CAACCCCTTTGGTCCCTCCCTTTGATGGGGCTCTGTTTCACTCTATTTCACCTCT 600
Db 7504 CAACCCCTTTGGTCCCTCCCTTTGATGGGGCTCTGTTTCACTCTATTTCACCTCT 7563
Qy 601 ATTAATCTTGCACACTG 617
Db 7564 ATTAATCTTGCACACTG 7580

RESULT 28
ABN97978
ID ABN97978 standard; DNA; 46340 BP.
XX AC ABN97978;
XX AC ABN97978;
XX 01-AUG-2002 (first entry)
XX Human retroviral sequence HI3.
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
KW

KW multiple sclerosis; ds.
XX Human endogenous retrovirus.
OS WO9867395-A1.
FN 29-DEC-1999.
XX 23-JUN-1999; 99WO-FR001513.
PD 23-JUN-1998; 98FR-00007920.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PR Alliel PM, Perin J, Rieger F;
XX WPI; 2000-160587/14.
DR New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
XX for diagnosis, treatment and prevention of autoimmune and neurological
XX diseases.
XX Claim 15; Page 186-199; 225pp; French.
XX The present invention relates to new nucleic acid sequences of human
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
XX Regulatory elements associated with HERV-7q may alter expression of other
XX genes (even remote genes) on the same chromosome, inducing immunological
XX and/or neurological changes (which may be pathological or protective/
XX curative). HERV-7q peptides can be used to improve efficiency of the
XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
XX sequences can be used in immunogenic or vaccinating compositions, for
XX protection against autoimmune diseases, particularly multiple sclerosis.
XX The peptides may also be used (by sequence comparison) to detect/identify
XX endogenous retroviruses that are abnormally expressed in cancer,
XX neuropathologies or other autoimmune diseases. The present sequence was
XX used to illustrate the invention
XX SQ Sequence 46340 BP; 16104 A; 8738 C; 8434 G; 13064 T; 0 U; 0 Other;
Query Match 80.4%; Score 510.6; DB 3; Length 46340;
Best Local Similarity 90.0%; Pred. No. 4.2e-147;
Matches 575; Conservative 0; Mismatches 49; Indels 15; Gaps 2;
QY 1 CCCTGTATCTTAACTCTTGTAAAGTTTGTCTTCCAGATCAAACTGTAAACTA 60
DB 33280 CCCTGTATCTTAACTCTTGTAAAGTTTGTCTTCCAGATCAAACTGTAAAGCTA 33339
QY 61 CAAATGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
DB 33340 CAAATGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 33399
QY 121 CTGACCGGCTGTAGCCCATGCTCCGATGTTAAATGACATTAAGAGCACCCCTCCCGAG 180
DB 33400 CTGACCGGCTGTAGCCCATGCTCCGATGTTAAATGACATTAAGAGCACCCCTCCCGAG 33459
QY 181 GAAATCTCAATGACACACCCCTACTATGCCAATTCAGCGGAGCAGTGTAGAGCGGT 240
DB 33460 GAAATCTCAATGACACACCCCTACTATGCCAATTCAGCGGAGCAGTGTAGAGCGGT 33519
QY 241 CATCAGCAACCTCCCAACAGCATTGGGTTTTTCTGTGTTGAGAGGGGGACTGAGAGAC 300
DB 33520 CGTCAGCAACCTCCCAACAGCATTGGGTTTTTCTGTGTTGAGTGGGGGACTGAGAGAC 33579
QY 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAGAATCCCTAAGCTAGCTGGGAGGTGACT 360
DB 33580 AGGATTAGCTGGATTTCTTAGGCCCAACGAAGAATCCCAAGCCCTAGCTGGGAGGTGACC 33639
QY 361 GCATCCACCTTAAACATGGGCTTGCACCTTAGCTCACCACCGACCAATC----- 411
DB 33640 ACATCCACCTTAAACATGGGCTTGCACCTTAGCTCACCACCGACCAATCAGGTAGTAA 33699
QY 412 AGAGAGCTCACTAAATGCTAAATAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATC 471

DB 33700 AGAGAGCTCACTAAATGCTAATTAGACAAAACAGGAGTAAAAAATAGCCAATCATC 33759
QY 472 TATTGCTGTAGACACAGCGGAGGACAAAGGATCGGATATTAACCCAGGCATTGAGC 531
DB 33760 TATCGCTGTAGACACAGCGGAGGACAAATGATCGGATATTAACCCAGGCATTCAAGC 33819
QY 532 CGGCACGGCAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTCTTCTCACTCTA 591
DB 33820 CGGCACGGCTACTCTTTTGGGTCCTCCCTTTGTATGGGAGCTCTCTCT-----GT 33873
QY 592 TTTCACTCTATTAAATCTTGCAACTGAAAAAATAAAAAA 630
DB 33874 CTTCACTCTATTAAATATTGCAACTGCAAAAAAATAA 33912
RESULT 29
ACN44334/C
ID ACN44334 standard; DNA; 161334 BP.
XX ACN44334;
AC ACN44334;
XX 18-NOV-2004 (first entry)
XX Human genomic sequence hCG32959.
DE Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
KW Homo sapiens.
OS WO2003073826-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 730; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: this patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published

XX SQ Sequence 161334 BP; 46474 A; 28689 C; 31549 G; 54602 T; 0 U; 20 Other;
Query Match 80.4%; Score 510.6; DB 11; Length 161334;
Best Local Similarity 90.0%; Pred. No. 7.7e-147;
Matches 575; Conservative 0; Mismatches 49; Indels 15; Gaps 2;
QY 1 CCCTGTATCTTAACTCTTGTAAAGTTTGTCTTCCAGATCAAACTGTAAACTA 60
DB 100611 CCCTGTATCTTAACTCTTGTAAAGTTTGTCTTCCAGATCAAACTGTAAAGCTA 100552

QY 61 CAAATGTTCTTCAAAATGGAGCACCAGATCGAGTCCATGACTAAGATCCACCGTGACCC 120
DB 100551 CAAATGTTCTTCAAAATGGAGCACCAGATCGAGTCCATGACTAAGATCCACCGTGACCC 100492
QY 121 CTGGACCGGCTCTAGCCCATCTCCGATGTTAATGACATGGAAGGACCCCTCCCGAG 180
DB 100491 CTGGACCGGCTCTAGCCCATCTCCGATGTTAATGACATGGAAGGACCCCTCCCGAG 100432
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGACGGT 240
DB 100431 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGACGGT 100372
QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 100371 CGTCAGCAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 100312
QY 301 AGGACTAGCTGATTTCTTAGGCCAAGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 100311 AGGACTAGCTGATTTCTTAGGCCAAGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 100252
QY 361 GCATCACCCTCTAAACATGGGGCTTGCAACTTAGCTCAGCCGACCAATC----- 411
DB 100251 ACATCCACCTTTAAACACTGGGCTTGCAACTTAGCTCAGCCGACCAATCAGGTAGTAA 100192
QY 412 AGAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATC 471
DB 100191 AGAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGCCCAATCATC 100132
QY 472 TATTGCTGAGAGCAGACGCGGAGGGAACAAGGATCGGGATATAAACCCAGGCAATTCGAGC 531
DB 100131 TATCGCTGAGAGCAGACGCGGAGGACATGATCGGATATAAACCCAGGCAATTCGAGC 100072
QY 532 CGGCAACGGCAACCCCTTTGGGTCCCTCTGTTGATGGGCTCTGTTTCACTCTA 591
DB 100071 CGGCAACGGCTACCTTTGGGTCCCTCTGTTGATGGGCTCTCTCT-----GT 100018
QY 592 TTTCACTCTAATAATCTGCAACTGAAAAAATAATA 630
DB 100017 CTTCACTCTAATAATCTGCAACTGAAAAAATAATA 99979

RESULT 30

ABN97929 standard; DNA; 10499 BP.

AC

ABN97929;

01-AUG-2002 (first entry)

Human retroviral sequence HERV-7q.

Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;

multiple sclerosis; ds.

Human endogenous retrovirus.

WO9967395-A1.

29-DEC-1999.

23-JUN-1999; 99WO-FR001513.

23-JUN-1998; 98FR-00007920.

(INRM) INSERM INST NAT SANTE & RECH MEDICALE.

Alliel PM, Perin J, Rieger F;

WPI; 2000-160587/14.

New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used

for diagnosis, treatment and prevention of autoimmune and neurological

PT

PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

Sequence 10499 BP; 3048 A; 2676 C; 2280 G; 2495 T; 0 U; 0 Other;

XX

Query Match 79.7%; Score 505.8; DB 3; Length 10499;

Best Local Similarity 91.1%; Pred. No. 6.3e-146;

Matches 564; Conservative 0; Mismatches 42; Indels 13; Gaps 2;

QY 1 CCCTGTATCTTTAAGCTCTTGTAAAGTTTGTCTCTCCAGAAATCAAAAAGTGTAAACCTA 60

DB 9274 CCCTGTATCTTTAAGCTCTTGTAAAGTTTGTCTCTCCAGAAATCAAAAAGTGTAAACCTA 9333

QY 61 CAAATGTTCTTCAAAATGGAGCACCAGATCGAGTCCATGACTAAGATCCACCGTGACCC 120

DB 9334 -----CAAATGAGGCCAAGATCGAGTCCAAAGCTAAGATCTACCCGACCC 9381

QY 121 CTGGACCGGCTCTAGCCCATCTCCGATGTTAATGACATTAAGGACCCCTCCCGAG 180

DB 9382 CTGGACCGGCTCTAGCCCATCTCCGATGTTAATGACATTAAGGACCCCTCCCGAG 9441

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGACGGT 240

DB 9442 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTTAGACGGT 9501

QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300

DB 9502 C-TGGGCCAAACCTCCCAACAGACACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 9560

QY 301 AGGACTAGCTGGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

DB 9561 AGGACTAGCTGGATTTCTTAGGCCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 9620

QY 361 GCATCCACCTCTAAACATCGGGCTTCGAACTTAGCTCAGACCCGACCAATCAGAGAGCTC 420

DB 9621 ACATCCACCTCTAAACATCGGGCTTCGAACTTAGCTCAGACCCGACCAATCAGAGAGCTC 9680

QY 421 ACTAAAATGCTAATAGGCAAAAATAGGAGTTAAGAAATAGCAATCATCTATTGCTG 480

DB 9681 ACTAAAATGCTAATAGGCAAAAATAGGAGTTAAGAAATAGCAATCATCTATTGCTG 9740

QY 481 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCCCGCAACCG 540

DB 9741 AGAGCAGCGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCCCGCAACCG 9800

QY 541 CAAACCCCTTTGGGTCCCTCTCCCTTTGATGGCGCTCTGTTTCACTCTATTCTCTCT 600

DB 9801 CAAACCCCTTTGGGTCCCTCTCCCTTTGATGGCGCTCTGTTTCACTCTATTCTCTCT 9860

QY 601 ATTAAATCTTGCAACTGAA 619

DB 9861 ATTAAATCTTGCAACTGCA 9879

RESULT 31

ACN44958

ID ACN44958 standard; DNA; 285020 BP.

XX

SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 Query Match 78.7%; Score 499.6; DB 4; Length 1894;
 Best Local Similarity 88.7%; Pred. No. 2.3e-144;
 Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCAAAATCTGTAATACTA 60
 DB 1217 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCGAAGCAGTAAATACTA 1276
 QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGAGTCCATGACTAAGATCCACCTGAGACC 120
 DB 1277 CAAATCTGTTCTTCAAAATGAGACACAGATGAGTCCATGAGTAAATCTACACGAGACC 1336
 QY 121 CTGACCGGCTCTAGCCATGCTCCGATGTTAAAGTAAATGACATTCGAAGCACCCTCCCGAG 180
 DB 1337 CTGACCGGCTCTAGCCATGCTCCGATGTTAAAGTAAATGACATTCGAAGCACCCTCCCGAG 1396
 QY 181 GAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCACGGGGAAGCAGTTAGAGCGGT 240
 DB 1397 GAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCACGAGGAGCAGTTAGAGTGT 1456
 QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
 DB 1457 TGTGGCCAACTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 1516
 QY 301 AGGACTAGCTGGATTTCTAGGCCCAAGAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 DB 1517 AGGAATAAATAGATTTCTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGGTGACC 1576
 QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTACTGCTCACACCCGACCAATC----- 411
 DB 1577 GCTTCCACCTTTAAACACCGGCTTGCAACTTACTGCTCACGCCCCCAACCAATCAGTACTAA 1636
 QY 412 AGAGAGCTCAATAATGCTAAATAGGCAAAAATAGAGGTAAAGAAATAGCCAAATCATC 471
 DB 1637 AGAGAGCTCAATAATGCTAAATAGGCAAAAATAGAGGTAAAGAAATAGCCAAATCATC 1696
 QY 472 TATTGCTGAGACACAGCGGAGGACAGGATCGGGATATAACCCAGGCATTCGAGC 531
 DB 1697 TGTGCTGACGACAGCAGGAGGGAATGATCGGGATATAACCCAGGCATTCGAGC 1756
 QY 532 CGGCAACCGCAACCCCTTTGGGTCCCTCTCTTGTATGGGCTCTGTTTCACTCTA 591
 DB 1757 CAGTACAGTACCTCTTTGGGTCCCTCTCTTGTATGGGCTCTGT----- 1806
 QY 592 TTTCACTCTATTAAATCTTGCAACTGCAAAATAGAAAATAGAAA 635
 DB 1807 CTTCACTCTATTAAATCTTGCAACTGCAAAATAGAAAATAGAAA 1850

RESULT 33
 ABA56337
 ID ABA56337 standard; DNA; 1894 BP.
 XX AC ABA56337;
 XX AC ABA56337;
 XX DT 01-FEB-2002 (first entry)
 XX DE Human foetal liver single exon nucleic acid probe #4642.
 XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX OS Homo sapiens.
 XX PN WO200157277-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US000669.
 XX PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 DR Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human foetal liver.
 PT Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human foetal liver. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 Query Match 78.7%; Score 499.6; DB 4; Length 1894;
 Best Local Similarity 88.7%; Pred. No. 2.3e-144;
 Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
 QY 1 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCAAAATCTGTAATACTA 60
 DB 1217 CCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTCTTCCAGAAATCGAAGCAGTAAATACTA 1276
 QY 61 CAAATTTGTTCTTCAAAATGAGACACAGATGAGTCCATGACTAAGATCCACCTGAGACC 120
 DB 1277 CAAATCTGTTCTTCAAAATGAGACACAGATGAGTCCATGAGTAAATCTACACGAGACC 1336
 QY 121 CTGACCGGCTCTAGCCATGCTCCGATGTTAAAGTAAATGACATTCGAAGCACCCTCCCGAG 180
 DB 1337 CTGACCGGCTCTAGCCATGCTCCGATGTTAAAGTAAATGACATTCGAAGCACCCTCCCGAG 1396
 QY 181 GAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCACGGGGAAGCAGTTAGAGCGGT 240
 DB 1397 GAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCACGAGGAGCAGTTAGAGTGT 1456
 QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
 DB 1457 TGTGGCCAACTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 1516
 QY 301 AGGACTAGCTGGATTTCTAGGCCCAAGAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 DB 1517 AGGAATAAATAGATTTCTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGGTGACC 1576
 QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTACTGCTCACACCCGACCAATC----- 411
 DB 1577 GCTTCCACCTTTAAACACCGGCTTGCAACTTACTGCTCACGCCCCCAACCAATCAGTACTAA 1636
 QY 412 AGAGAGCTCAATAATGCTAAATAGGCAAAAATAGAGGTAAAGAAATAGCCAAATCATC 471
 DB 1637 AGAGAGCTCAATAATGCTAAATAGGCAAAAATAGAGGTAAAGAAATAGCCAAATCATC 1696
 QY 472 TATTGCTGAGACACAGCGGAGGACAGGATCGGGATATAACCCAGGCATTCGAGC 531
 DB 1697 TGTGCTGACGACAGCAGGAGGGAATGATCGGGATATAACCCAGGCATTCGAGC 1756
 QY 532 CGGCAACCGCAACCCCTTTGGGTCCCTCTCTTGTATGGGCTCTGTTTCACTCTA 591
 DB 1757 CAGTACAGTACCTCTTTGGGTCCCTCTCTTGTATGGGCTCTGT----- 1806
 QY 592 TTTCACTCTATTAAATCTTGCAACTGCAAAATAGAAAATAGAAA 635

Fri Feb 25 16:26:30 2005

us-09-319-156b-6.rng

```
||||| 1807 CTTCACTCTATTAAATCTTGCAACTGCAAAATATAAAATAGAAA 1850
Db
RESULT 34
AAI35980
ID AAI35980 standard; DNA; 1894 BP.
AC AAI35980;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #4666 used to measure gene expression in human placenta sample.
DE
DE Probe #4666 used to measure gene expression in human placenta sample;
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
PN
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
PT
XX Claim 25; SEQ ID NO 4666; 654pp; English.
PS
XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
XX SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match 78.7%; Score 499.6; DB 4; Length 1894;
Best Local Similarity 88.7%; Pred. No. 2.3e-144;
Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;
Qy 1 CCCTGTATCTTAACTCCCTCTGTAACTTGTCTCTCCAGATCAAACTGTAATAACTA 60
Db 1217 CCCTGTATCTTAACTCCCTCTGTAACTTGTCTCTCCAGATCAAACTGTAATAACTA 1276
Qy 61 CAAATTTCTTCAATGGAGCACGACATGAGTCCATGATCAAGATCCACCGTGACCC 120
Db 1277 CAAATCTTCTTCAATGGAGCCCGACATGATGATGATGATGATGATGATGATGATGAT 1336
Qy 121 CTGACCGCGCTGTAGCCCATGCTCCGATGTTAATGACATTTAGGACACCCCTCCCGAG 180
Db 1337 CTGACCGCGCTGTAGCCCATGCTCTGATGTTAATGACATCAAAAGCACCCCTCCCGAG 1396
Qy 181 GAAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGAAGCAGTTAGAGCGGT 240
Db 1397 GAAATCTCACTGACACACCTCTACTAGCCCCCAATTCAGCAGGAAGCAGTTAGAGTGT 1456
Qy 241 CATCAGCCAACTCTCCCAACAGACACTTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
Db
1457 TGTTCGCCAACCTCCCAACAGCAGTTGGGTTTCTCTGTTGAGAGGGGGACTGAGAGAC 1516
Qy 301 AGGACTAGCTGGATTTCTTAGCCACGAGAGATCCCTAAGCTAGCTGGGAAGGTGACT 360
Db 1517 AGGAATAAATAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAAGGTGACC 1576
Qy 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 411
Db 1577 GCTTCACCTTTAAACACCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAA 1636
Qy 412 AGAGAGCTCACTAAATGCTTAATTAGGCAAAATAGAGGTAAGAAATAGCCCAATCATC 471
Db 1637 AGAGAGCTCACTAAATGCTTAATTAGGCAAAATAGAGGTAAGAAATAGCCCAATCATC 1696
Qy 472 TATTGCTGAGACACAGCGGGAGGACAGGATCGGATATATAACCCAGGCAATTCGAGC 531
Db 1697 TGTTCCTGACGACAGCAGGAGGACAAATGATCGGATATATAACCCAGGCAATTCGAGC 1756
Qy 532 CGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGATGGGCGTCTGTTTTCACCTCTA 591
Db 1757 CAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGT----- 1806
Qy 592 TTTCACCTCTAATAATCTTGCAACTGAAATAAAAAAAAAAAAAA 635
Db 1807 CTTCACTCTAATAATCTTGCAACTGCAAAATAAAAAAAAAAGAAA 1850
RESULT 35
ABA45822
ID ABA45822 standard; DNA; 1894 BP.
XX
XX ABA45822;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #4517.
DE
DE Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157271-A2.
PN
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
XX Claim 1; SEQ ID NO 4517; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
```

CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 78.7%; Score 499.6; DB 4; Length 1894;
Best Local Similarity 88.7%; Pred. No. 2.3e-144;
Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;

QY 1 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
DB 1217 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 1276
QY 61 CAAATTTCTTCAATGGAGACCCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 1277 CAAATCTCTTCAATGGAGACCCAGATGAGTCCATGAGTAAATCTACCGACCC 1336
QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGAGCACCCCTCCGAG 180
DB 1337 CTGACCGGCTCTAGCCCATGCTCTGATGTTTAATGACATTAAGAGCACCCCTCCGAG 1396
QY 181 GAAATCTCAATGCACACCCCTACTATGATGCTCCCAATTCAGCGGAGCAGTTAGACGGT 240
DB 1397 GAAATCTCAATGCACACCCCTACTATGATGCTCCCAATTCAGCGGAGCAGTTAGACGGT 1456
QY 241 CATCAGCACCTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
DB 1457 TGTGGCCAACTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 1516
QY 301 AGGACTAGCTGGATTTCTTAGGCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 1517 AGGATTAACCTAGATTTCTTAGACCACTAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 1576
QY 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATC----- 411
DB 1577 GCTTCCACCTTTAAACACCGGGCTTGCACTTAGCTCACACCCGACCAATCAGATATAA 1636
QY 412 AGAGAGCTCACTAAATGCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATC 471
DB 1637 AGAGAGCTCACTAAATGCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATC 1696
QY 472 TATTGCTGAGAGCAGCAGCGGGAGGCAAGGATCGGGATATAAACCCAGGCAATTCGAGC 531
DB 1697 TGTTCCTGACAGCAGCAGGAGGACATGATCGGATATAAACCCAGGCAATTCGAGC 1756
QY 532 CGGCAACGGCAACCCCTTTGGTCCCTCCCTTTGATGAGGGCCCTCTGTTTCACTCTA 591
DB 1757 CAGCTACAGCTACCCCTCTTTGGTCCCTCCCTTTGATGAGGGACTCTGT----- 1806
QY 592 TTTTCACTTATTAATCTTCACTGAAATGAAAAAAGAAAAA 635
DB 1807 CTTTCACTTATTAATCTTCACTGAAATGAAAAAAGAAAA 1850

RESULT 36

ABA25978 standard; DNA; 1894 BP.

XX ABA25978;

XX ABA25978;

XX 23-JAN-2002 (first entry)

DT

XX DE Probe #4444 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; hs.

OS Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000666.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human hearts.

XX Claim 1; SEQ ID NO 4444; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 78.7%; Score 499.6; DB 4; Length 1894;

Best Local Similarity 88.7%; Pred. No. 2.3e-144;

Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;

QY 1 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
DB 1217 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 1276
QY 61 CAAATTTCTTCAATGGAGACCCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 1277 CAAATCTCTTCAATGGAGACCCAGATGAGTCCATGAGTAAATCTACCGACCC 1336
QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTAAGAGCACCCCTCCGAG 180
DB 1337 CTGACCGGCTCTAGCCCATGCTCTGATGTTTAATGACATTAAGAGCACCCCTCCGAG 1396
QY 181 GAAATCTCAATGCACACCCCTACTATGATGCTCCCAATTCAGCGGAGCAGTTAGACGGT 240
DB 1397 GAAATCTCAATGCACACCCCTACTATGATGCTCCCAATTCAGCGGAGCAGTTAGACGGT 1456
QY 241 CATCAGCACCTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
DB 1457 TGTGGCCAACTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGGGACTGAGAGAC 1516
QY 301 AGGACTAGCTGGATTTCTTAGGCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360

Query Match		78.7%;	Score 499.6;	DB 4;	Length 1894;					
Best Local Similarity		88.7%;	Pred. No. 2.3e-144;							
Matches 571;		Conservative	0;	Mismatches 54;	Indels 19; Gaps 2;					
Qy	1	CCCTGTATCTTTAACTCTCTTTGTTTAAAGTTTGTCTCTTCCAGAAATCAAAATCTGTAATACTA	60							
Db	1217	CCCTGTATCTTTAACTCTCTTTGTTTAAAGTTTGTCTCTTCCAGAAATCGAAGCAGTAAATACTA	1276							
Qy	61	CAAAATCTTCTTCAAAATGGAGCACACAGATGCGAGTCCATGACTAAGATCCACCGTGGACCC	120							
Db	1277	CAAAATCTTCTTCAAAATGGAGCCCCAGATGCGAGTCCATGAGTAAATCTTACCACGGACCC	1336							
Qy	121	CTGGACGGGCTGCTGACCCCATGCTCCGATGTTTAATGACATTTGAAGCACACCCCTCCCGAG	180							
Db	1337	CTGGACGGGCTGCTGACCCCATGCTCTGATGTTTAATGACATCAAAAGCACCCCTCCCGAG	1396							
Qy	181	GAATCTCAATGTCACAAACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT	240							
Db	1397	GAATCTCAATGTCACAAACCCCTACTATGCCCCCAATTCAGCAGAGACGATTAGAGTGGT	1456							
Qy	241	CATCAGCAACCTCCCAACAGCACTTTGGGTTTTTCTGTTGAGAGGGGGGACTGAGAGAC	300							
Db	1457	TGTTGGCAACCTCCCAACAGCAGTTGGGTTTTTCTGTTGAGAGGGGGGACTGAGAGAC	1516							
Qy	301	AGGACTAGCTGGATTTCTTAGGCCCAACGAGAGATCCCTAAGCCTAGCTGGGAGGTGACT	360							
Db	1517	AGGAATAACTAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAGGTGACC	1576							
Qy	361	GCATCCACCTTAACATGATGGGCTTGCAACTTAGCTCACACCCGACCAATC-----	411							
Db	1577	GCATCCACCTTTAAACACCCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAA	1636							
Qy	412	AGAGAGCTCACTAAATGCTAAATTAGGCACAAATAGGAGGTAAAGAAATAGCCAAATCATC	471							
Db	1637	AGAGAGCTCACTAAATGCTAAATTAGGCACAAATAGGAGGTAAAGAAATAGCCAAATCATC	1696							
Qy	472	TATTGCTGTAGAGCACAGCGGGAGGACAGAGATCGGGATATAAACCCGAGCATTCGAGC	531							
Db	1697	TATTGCTGTAGAGCACAGCGGGAGGACAGATGATCGGGATATAAACCCGAGCATTCGAGC	1756							
Qy	532	CGGCAACGGCAACCCCTTTGGGTCCTTGGGTCCTTGGTATGGGCGCTCTGTTTCACTCTA	591							
Db	1757	CAGCTACAGCTACCCCTTTGGGTCCTTGGGTCCTTGGTATGGGCGCTCTGTTTCACTCTA	1806							
Qy	592	TTTCACCTCTATTAAATCTTGCAACTGCAAAAAAATAAAAAAATAAAAAA 635								
Db	1807	CTTCACCTCTATTAAATCTTGCAACTGCAAAAAAATAAAAAAATAAAAAA 1850								
RESULT 38										
AAK04516										
ID AAK04516 standard; DNA; 1894 Bp.										
XX AC AAK04516;										
XX DT 05-NOV-2001 (first entry)										
XX DE Human brain expressed single exon probe SEQ ID NO: 4507.										
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;										
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;										
XX OS Homo sapiens.										
XX PN WO200157275-A2.										
XX PD 09-AUG-2001.										
XX PF 30-JAN-2001; 2001WO-US000667.										
XX PR 04-FEB-2000; 2000US-0180312P.										
XX PR 26-MAY-2000; 2000US-0207456P.										
XX PR 30-JUN-2000; 2000US-00608408.										
XX PR 03-AUG-2000; 2000US-00632366.										
XX PR 21-SEP-2000; 2000US-0234687P.										
XX PR 27-SEP-2000; 2000US-0236359P.										
XX PR 04-OCT-2000; 2000GB-00024263.										
(MOLE-) MOLECULAR DYNAMICS INC.										
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;										
XX PF WPI; 2001-488900/53.										
XX KW Human genome-derived single exon nucleic acid probes useful for analyzing										
XX KW gene expression in human bone marrow.										
XX PS Example 4; SEQ ID NO 4575; 658pp + Sequence Listing; English.										
XX CC The present invention provides a number of single exon nucleic acid										
XX CC probes which are derived from genomic sequences expressed in the human										
XX CC bone marrow. They can be used to measure gene expression in bone marrow										
XX CC samples, which may enable the improved diagnosis and treatment of cancers										
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of										
XX CC the probes of the invention										
XX PR Sequence 1894 Bp; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;										
XX SQ										

Db	1517	AGGAATAACTAGATTTCTTAGACCAACTAAGAAATCCCTAAGACTAGCTGGGAGGTGACC	1576
Qy	361	GCATCCACCTCTAAACATGGGGCTTTGCAACTTAGCTCACACCCGACCAATC-----	411
Db	1577	GTTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTCAGGCCAACCAATCAGATACTAA	1636
Qy	412	AGAGAGCTCACTAAATGCTAATTAGGCGMAAATAGGAGGTAAAGAAATAGCCAAATCATC	471
Db	1637	AGAGAGCTCACTAAATGCTAATTAGGCGMAAATAGGAGGTAAAGAAATAGCCAAATCATC	1696
Qy	472	TATTGCTGTAGAGCACAGCGGGAGGACAGAGATCGGGATATAAACCCGAGCATTCGAGC	531
Db	1697	TATTGCTGTAGAGCACAGCGGGAGGACAGATGATCGGGATATAAACCCGAGCATTCGAGC	1756
Qy	532	CGGCAACGGCAACCCCTTTGGGTCCTTGGGTCCTTGGTATGGGCGCTCTGTTTCACTCTA	591
Db	1757	CAGCTACAGCTACCCCTTTGGGTCCTTGGGTCCTTGGTATGGGCGCTCTGTTTCACTCTA	1806
Qy	592	TTTCACTCTATTAAATCTTGCAACTGCAAAAAAATAAAAAAATAAAAAA 635	
Db	1807	CTTCACTCTATTAAATCTTGCAACTGCAAAAAAATAAAAAAATAAAAAA 1850	
RESULT 37			
AAK30018			
ID AAK30018 standard; DNA; 1894 Bp.			
XX AC AAK30018;			
XX DT 06-NOV-2001 (first entry)			
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 4575.			
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;			
XX KW microarray; cancer; leukemia; lymphoma; myeloma; ss.			
XX OS Homo sapiens.			
XX PN WO200157276-A2.			
XX PD 09-AUG-2001.			
XX PF 30-JAN-2001; 2001WO-US000668.			
XX PR 04-FEB-2000; 2000US-0180312P.			
XX PR 26-MAY-2000; 2000US-0207456P.			
XX PR 30-JUN-2000; 2000US-00608408.			
XX PR 03-AUG-2000; 2000US-00632366.			
XX PR 21-SEP-2000; 2000US-0234687P.			
XX PR 27-SEP-2000; 2000US-0236359P.			
XX PR 04-OCT-2000; 2000GB-00024263.			
(MOLE-) MOLECULAR DYNAMICS INC.			
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;			
XX PF WPI; 2001-488900/53.			
XX KW Human genome-derived single exon nucleic acid probes useful for analyzing			
XX KW gene expression in human bone marrow.			
XX PS Example 4; SEQ ID NO 4575; 658pp + Sequence Listing; English.			
XX CC The present invention provides a number of single exon nucleic acid			
XX CC probes which are derived from genomic sequences expressed in the human			
XX CC bone marrow. They can be used to measure gene expression in bone marrow			
XX CC samples, which may enable the improved diagnosis and treatment of cancers			
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of			
XX CC the probes of the invention			
XX PR Sequence 1894 Bp; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;			
XX SQ			

Db 1807 CTTCACTCTATTAAATCTTTGCAACTGCAAAAATAGAAA 1850

RESULT 39
ABS29670
ID ABS29670 standard; DNA; 1894 BP.
XX
AC ABS29670;
XX
XX 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID NO 4660.
XX
XX Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000664.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-00608408.
XX
XX PR 03-AUG-2000; 2000US-00632366.
XX
XX PR 21-SEP-2000; 2000US-0234687P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX PI WPI; 2001-488998/53.
XX
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human adult liver.
XX
XX PS Claim 1; SEQ ID NO 4660; 658pp; English.
XX
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (I) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
XX CC involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX CC associated with coronary heart disease. ABS25011-AB951005 represent human
XX CC liver single exon nucleic acid probes of the invention. Note: The
XX CC sequence information for this patent does not appear in the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1894 BP; 557 A; 500 G; 476 T; 0 U; 0 Other;

Query Match 78.7%; Score 499.6; DB 4; Length 1894;
Best Local Similarity 88.7%; Pred. No. 2.3e-144;
Matches 571; Conservative 0; Mismatches 54; Indels 19; Gaps 2;

Qy 1 CCCTGTATCTTTAACTCCCTGTGTAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
Db 1217 CCCTGTATCTTTAACTCCCTGTGTAGTTTGTCTTTTCCAGATCGAAGCAGTAAACTA 1216

Qy 61 CAAATTGTTCTTCAAATGGAGCACCAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 1277 CAAATCGTTCTTCAAATGGAGCCCGCAGATGCATGCTCAATGATGATAAATCTACCACGGACCC 1336

QY	121	CTGCACGGGCTCTAGCCCATGCTCCGATGTTAAATGACATTGAAGGCACCCCTCCCGAG	180
Db	1337	CTGCACGGGCTCTAGCCCATGCTCTGATGTTAAATGACATCAAGAGCACCCCTCCCGAG	1396
QY	181	GAATCTCAACTCGACACACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGACGGT	240
Db	1397	GAATCTCAACTCGACACACCTCTACTACGCCCCAAATTCAGCAGGAAGCAGTTAGATGGT	1456
QY	241	CATCAGCCAACTCTCCCAACAGCACTTGGGTTTTCTGTTCAGAGGGGGACTTGAGAGAC	300
Db	1457	TGTTGGCCAACTCCCCAACAGCAGTTGGGTTTTCTGTTCAGAGGGGGACTTGAGAGAC	1516
QY	301	AGACTAGCTGGATTTCTTAGGCCAAGAGAAATCCCTAAGCCCTAGCTGCTGGGAAGGTGACT	360
Db	1517	AGGAATAACTAGATTTCTCTAGACCAACTAAGAAATCCCTAAGACTAGCTGCTGGGAAGGTGACC	1576
QY	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC-----	411
Db	1577	GCITCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAA	1636
QY	412	AGAGAGCTCACTAAATGCTAAATTAGCCAAAATAGGAGCTAAGAAATAGCCAATCATC	471
Db	1637	AGAGAGCTCACTAAATGCTAAATTAGGCCAAAACAGGAGATAAAGAATTAGCCAAATCATC	1696
QY	472	TATTGCTTGAGAGCACACGCGGGAGGACAAGGATCGGGATATAAACCCAGGCATTCGAGC	531
Db	1697	TGTTGCTTGACAGCACACGAGGAGGACAATGATCGGGATATAAACCAGGCATTCGAGC	1756
QY	532	CGGCAACGGCAACCCCTTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTTCACTCTA	591
Db	1757	CAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGT-----	1806
QY	592	TTTCACCTCTATAAATCTTGCAACTGAAAAAATAAAAAAAAAA	635
Db	1807	CTTTACCTCTATAAATCTTGCAACTGCAAAAATAAAAAATAGAA	1850

RESULT 40

RESULI 4
AAT04422

AA104422
ID AA104422 standard: DNA: 1894 BP.

ID
YY
AA10

AC
X

AC YY AA100

XX 09-01

D.T. 09-01
v.v.

XX DE 1964

DE
vv
Prob

XX
XX
XX

KW	Prob.
1	0.1
2	0.2
3	0.3
4	0.2
5	0.1
6	0.1

KW
inf l.
yy

XX 5 17.04.0

SO
Homo

XX 1300

PN WO20

XX 3

PD 09-A

XX

PF 29-J

XX

PR 04-F

PR 26-M

PR 30-J

PR 03-A

PR 21-S

PR 27-S

PR 04-0

PT a human breast.

PI a human breast.
yy

Claim 25; SEQ ID NO 4413; 322pp; English.

The present invention relates to novel single exon nucleic acid probes. The present sequence is one such probe. The probes are useful for measuring human gene expression in a human breast sample, where the probes hybridizes at high stringency to a nucleic acid expressed in the human breast. The probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic aetiology. The diseases include: breast cancer, disorders of development, inflammatory diseases of the breast, fibrocystic changes, proliferative breast disease and non-carcinoma tumours. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form as part of a file transfer via info@publinter.com published pct sequences

XX
CO
commence 1894 BP. 557 A: 500 C: 361 G: 476 T: 0 U: 0 Other;

Country	Match	78	79	Score	499.6	DB	5	Length	1894
---------	-------	----	----	-------	-------	----	---	--------	------

Query Match

Best Local Similarity 88.7%; Fred. NO: 2.3e-144; Indels 19;
Mismatch 0; Mismatches 54;

QY	1	CCCTGTATCTTTAAACCTCCCTGTTTAAAGTTTGTCTCTTCAGAAATCAAAA	CTGTAAAACTA	60
DB	1217	CCCTGTATCTTTAAACCTCCCTGTTTAAAGTTTGTCTCTTCAGAAATCAAAA	CTGTAAAACTA	1276
QY	61	CABAATTGTTCTTTCAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC		120
DB	1277	CABAATCGTTCTTTCAAATGGAGCCCCAGATGCAGTCCATGAGTAAATCTACCA	CGGACCC	1336
QY	121	CTGGACCGGCGCTGCTAGGCCCATGCTCCGATGTTTAAATGACATTTGAAGGCA	CCCCCTCCCGAG	180
DB	1337	CTGGACCGGCGCTGCTAGGCCCATGCTCGATGTTTAAATGACATCAAAAGC	ACCCCTCCCGAG	1396
QY	181	GAATCTCAATGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTTAGAGCGGT		240
DB	1397	GAATCTCAATGCACAAACCTCTACTAGCCCCCAATTCAGCAGAGACGATTTAGATGGT		1456
QY	241	CATCAGCCAAACCTCCCCAAACAGCACTTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC		300
DB	1457	TGTTGGCCAAACCTCCCCAAACAGCAGTTGGGTTTTCTGTTGAGAGGGGGACTGAGAGAC		1516
QY	301	AGGACTAGCTGGAATTTCTTAGGGCCAAACGAAGAATCCCTAAGCCCTAGCTGGGAAGTGACT		360
DB	1517	AGCAATAACTAGATTTCTTAGACCAACCTAAGAATCCCTAAGACTAGCTGGGAAGTGACC		1576
QY	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC	-----	411
DB	1577	GCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACGCCCAACCAATCAGATACTAA		1636
QY	412	AGAGAGCTCACTPAAAAATGCTAATTAGGCCAAAATAAGGAGGTAAAGAAATAGCCAATCATC		471
DB	1637	AGAGAGCTCACTPAAAAATGCTAATTAGGCCAAAACAGGAGATAAAGAAATAGCCAATCATC		1699
QY	472	TAATTCCTGAGACACAGCGGAGGAGCAAGGATCGGATATAAACCCACGAGCATTCGAGC		531
DB	1697	TGTTGCTTGACACAGCAGGAGGGACATGATCGGATATAAACCCACGAGCATTCGAGC		1759
QY	532	CGGCAACGGCAACCCCTTTTGGGTCCTCCCTTGTATGGGCGCTCTGTTTTCACTCTA		591
DB	1757	CAGCTACAGCTACCCCTCTTTGGGTCCTCCCTTGTATGGAGCTCTGT	-----	1809
QY	592	TTTCACTCTATTAAATCTTGCAACTGAAAAA	AAAAAAAAAAAAA	635
DB	1807	CTTCACTCTATTAAATCTTGCAACTGCAAAAA	TAAAAATAGAAA	1850

Research completed: February 27, 2005. 07:45:38

Search completed: FEBRU
Job time : 339 546 SECS

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:54:34 ; Search time 111.885 Seconds
(without alignments)
9286.612 Million cell updates/sec

Title: US-09-319-156B-6
Perfect score: 635
Sequence: 1 cccgtatcttcaactctct.....tgaaaaaiaaaaaaaaaa 635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	635	100.0	635	4	US-08-979-847B-102
2	545	85.8	1329	4	US-08-979-847B-108
3	533.8	84.1	2946	3	US-09-175-928-3
4	516.2	81.3	13537	4	US-09-949-016-17382
5	515.8	81.2	2763	4	US-09-949-016-5640
6	510.6	80.4	145320	4	US-09-949-016-15858
7	453	71.3	99580	4	US-09-949-016-17411
8	448	70.5	168394	4	US-09-949-016-13002
9	442.6	69.7	77772	4	US-09-949-016-17417
10	442.6	69.7	77997	4	US-09-949-016-12249
11	405.4	63.8	601	4	US-09-949-016-44300
12	367.2	57.8	149971	4	US-09-949-016-13590
13	292	46.0	245286	4	US-09-949-016-15497
14	291.6	45.9	601	4	US-09-949-016-133739
15	288	45.4	194790	4	US-09-949-016-15393
16	273.2	43.0	601	4	US-09-949-016-44296
17	269	42.4	219964	4	US-09-949-016-15086
18	262.8	41.4	780	4	US-09-573-080A-385
19	252.2	39.7	601	4	US-09-949-016-103638
20	252.2	39.7	601	4	US-09-949-016-103639
21	251	39.5	601	4	US-09-949-016-103637
22	251	39.5	256287	4	US-09-949-016-14608
23	249.4	39.3	601	4	US-09-949-016-44299
24	245.8	38.7	84571	4	US-09-949-016-17420
25	245.6	38.7	140224	4	US-09-949-016-17002
26	239.4	37.7	89584	4	US-09-949-016-17068
27	233.6	36.8	279	1	US-08-686-878A-50

28	233.6	36.8	279	1	US-08-721-489-4	Sequence 4, Appli
29	219.6	34.6	8523	4	US-09-573-080A-21	Sequence 21, Appl
30	196.6	31.0	456	4	US-09-621-976-9366	Sequence 9366, Ap
31	195.6	30.8	601	4	US-09-949-016-128750	Sequence 128750,
32	180.4	28.4	154023	4	US-09-949-016-17057	Sequence 17057, A
33	178	28.0	601	4	US-09-949-016-44301	Sequence 44301, A
34	176.8	27.8	251672	4	US-09-949-016-195792	Sequence 195792,
35	176.8	27.8	251682	4	US-09-949-016-117296	Sequence 17296, A
36	176.8	27.8	251682	4	US-09-949-016-11973	Sequence 11973, A
37	175	27.6	601	4	US-09-949-016-44310	Sequence 44310, A
38	174.6	27.5	116652	4	US-09-949-016-13413	Sequence 13413, A
39	174	27.4	39686	4	US-09-949-016-13633	Sequence 13633, A
40	174	27.4	49487	4	US-09-949-016-11770	Sequence 11770, A
41	166.8	26.3	131631	4	US-09-949-016-11757	Sequence 11757, A
42	164.8	26.0	57507	4	US-09-949-016-15019	Sequence 15019, A
43	156	24.6	64291	4	US-09-949-016-16278	Sequence 16278, A
44	156	24.6	117410	4	US-09-949-016-12262	Sequence 12262, A
45	147.2	23.2	601	4	US-09-949-016-202370	Sequence 202370,

ALIGNMENTS

RESULT 1

US-08-979-847B-102
; Sequence 102, Application US/08979847B
; Patent No. 6582703

GENERAL INFORMATION:

APPLICANT: PERRON, HERVE
BESEME, FREDERIC
BEDIN, FREDERIC
PARANHOS-BACCALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUREK, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE, PLC

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER:

APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-No. 6582703-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BERRIDGE, WILLIAM P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 39046A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:

LENGTH: 635 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 102:

US-08-979-847B-102

Query Match 100.0%; Score 635; DB 4; Length 635;
Best Local Similarity 100.0%; Pred. No. 2e-203;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60

QY 61 CAAATGTTCTTCAATGAGACACAGATGGAGTCCATGACTAAAGTCCACGTTGACCC 120
DB 61 CAAATGTTCTTCAATGAGACACAGATGGAGTCCATGACTAAAGTCCACGTTGACCC 120

QY 121 CTGGACCGGCTGTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 121 CTGGACCGGCTGTAGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG 180

QY 181 GAAATCTCAACTGCACAAACCTCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240
DB 181 GAAATCTCAACTGCACAAACCTCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240

QY 241 CATAGCCAACTCTCCCAACAGACCTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 241 CATAGCCAACTCTCCCAACAGACCTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300

QY 301 AGGACTAGCTGGATTTCTTAGCCACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 301 AGGACTAGCTGGATTTCTTAGCCACAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360

QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
DB 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCCCAATCATCTATTGCCTG 480
DB 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGAAATAGCCCAATCATCTATTGCCTG 480

QY 481 AGAGCACAGGGAGGACAGGATCGGATATAAACCCAGGATTCAGAGCGGCAACGG 540
DB 481 AGAGCACAGGGAGGACAGGATCGGATATAAACCCAGGATTCAGAGCGGCAACGG 540

QY 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTCACTCT 600
DB 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTCACTCT 600

QY 601 ATTAATCTTGCACTGAAAAA 635
DB 601 ATTAATCTTGCACTGAAAAA 635

RESULT 2
US-08-979-847B-108
; Sequence 108, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUXE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
THERAPEUTIC PURPOSES
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA

COUNTRY: USA
ZIP: 22320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,847B
FILING DATE: 26-No. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-08-979-847B-108

Query Match 85.8%; Score 545; DB 4; Length 1329;
Best Local Similarity 92.4%; Pred. No. 7.2e-173;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 CCCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60
DB 257 CCCTGTATCTTCAACTCTTGTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAAGT 316

QY 61 CAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAAGTCCACGTTGACCC 120
DB 317 CAAATGTTCTTCAAAATGGAAACCCAGATGAGTCCATGACTAAATCTACCGTGGACCC 376

QY 121 CTGACCGGCTGTAGCCATGCTCCATGTTAATGACATTAAGGACACCCCTCCCGAG 180
DB 377 CTGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTAAGTCAACCGCTCCCGAG 436

QY 181 GAAATCTCAACTGCACAAACCTCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240
DB 437 GAAATCTCAACTGCACAAACCTCTACTACATCCCAATTCAGTAGGAGCAGTTAGACGAGT 496

QY 241 CATCAGCAACCTCCCAACAGACCTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
DB 497 TGTGAGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGAGGGTGGACTGAGAGAC 556

QY 301 AGGACTAGCTGATTTCTTAGCCCAACCAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
DB 557 AGGACTAGCTGATTTCTTAGGCTGACTAAAGAAATCCCAAGCTTANCTGGGAAGGTGACC 616

QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
DB 617 GCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 676

QY 421 ACTAAATGCTAATTAGCAAAAAATAGAGGTAAGAAATAGCCCAATCATCTATTGCCTG 480
DB 677 ACTAAATGCTAATCAGGCAAAAAACAGGAGGTAAAGCAATAGCCCAATCATCTATTGCCTG 736

QY 481 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCAGAGCGGCAACGG 540
DB 737 AGAGCACAGCGGAGGACAGGATCGGATATAAACCCAGGATTCAGAGCGGCAACAG 796

QY 541 CAAACCCCTTTGGGTCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTCACTCT 600
DB 797 CAAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTCACTCTATTCACTCT 856

QY 601 ATTAATCTTGCAACTGAA 619

Db 857 ATTAATCATGCAACTGCA 875
|||||

RESULT 3

US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.AJ172A
; CURRENT APPLICATION NUMBER: US/09/175,928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3

Query Match 84.1%; Score 533.8; DB 3; Length 2946;
Best Local Similarity 91.5%; Pred. No. 7.2e-169;
Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTGTTCTTCCAGAAATCAAAACTGTAATAACTA 60
Db 2323 CCTGTATCTTTAACTCTTGTAAAGTTGTTCTTCCAGAAATCAAAACTGTAATAACTA 2382

QY 61 CAAATTGTTCTTCAATGGAGCCAGATGAGTCCATGACTAAGATCCACACCTGACACC 120
Db 2383 -----CAAAATGGAGCCAGATGAGTCCAAAGATTAAGATCTACCCGACACC 2430

QY 121 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
Db 2431 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 2490

QY 181 GAAATCTCAAATCGACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 2491 GAAATCTCAGCTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 2550

QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 300
Db 2551 CGTCGGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGATGGGGAAGCTGAGAGAC 2610

QY 301 AGGACTAGCTGGATTCTCTAGGCAACGAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 2611 AGGACTAGCTGGATTCTCTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACC 2670

QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACAACCCGACCAATCAGAGAGCTC 420
Db 2671 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACAACCTGACCAATCAGAGAGCTC 2730

QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCGCTG 480
Db 2731 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCGCTG 2790

QY 481 AGAGCAGCGGGAGGACAAAGATCGGGATATAAACCCAGGCAATTCAGCCGCGCAACGG 540
Db 2791 AGAGCAGCGGGAGGACAAATGATCGGGATATAAACCCAGGCTTCGAGCCGCGCAACGG 2850

QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 2851 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 2910

QY 601 ATTAATCTTGCACACTGAAAAAAGAAAAAAGAAAAA 635
Db 2911 ATTAATCTTGCACACTGCAAAAAAAGAAAAAAGAAAAA 2945

RESULT 4

US-09-949-016-17382
; Sequence 17382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17382
; LENGTH: 13537
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17382

Query Match 81.3%; Score 516.2; DB 4; Length 13537;
Best Local Similarity 91.1%; Pred. No. 1.7e-162;
Matches 564; Conservative 0; Mismatches 43; Indels 12; Gaps 1;

QY 1 CCTGTATCTTTAACTCTTGTAAAGTTGTTCTTCCAGAAATCAAAACTGTAATAACTA 60
Db 10931 CCTGTATCTTTAACTCTTGTAAAGTTGTTCTTCCAGAAATCAAAACTGTAATAACTA 10990

QY 61 CAAATTGTTCTTCAATGGAGCCAGATGAGTCCATGACTAAGATCCACACCTGACACC 120
Db 10991 -----CAAAATGGAGCCAGATGAGTCCAAAGATTAAGATCTACCCGACACC 11038

QY 121 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 180
Db 11039 CTGGACCGGCTGCTAGCCCATGCTCGATGTTTAATGACATTAAGGACACCCCTCCCGAG 11098

QY 181 GAAATCTCAAATCGACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 11099 GAAATCTCAGCTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 11158

QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGACCTGAGAGAC 300
Db 11159 CGTCGGCAACCTCCCAACAGCACTTGGGTTTCTGTTGAGATGGGGAAGCTGAGAGAC 11218

QY 301 AGGACTAGCTGGATTCTCTAGGCAACGAGAAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
Db 11219 AGGACTAGCTGGATTCTCTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACC 11278

QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACAACCCGACCAATCAGAGAGCTC 420
Db 11279 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACAACCTGACCAATCAGAGAGCTC 11338

QY 421 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCGCTG 480
Db 11339 ACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCGCTG 11398

QY 481 AGAGCAGCGGGAGGACAAAGATCGGGATATAAACCCAGGCAATTCGAGCCGCGCAACGG 540
Db 11399 AGAGCAGCGGGAGGACAAATGATCGGGATATAAACCCAGGCTTCGAGCCGCGCAACGG 11458

QY 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
|||||

Db	11459	CAACCCCTTTGGGTCCTCCCTTTGTATGGAGCTCTGTTTCATGCTATTTCACCTCT	11518
Qy	601	ATTAATCTTGCAACTGAA	619
Db	11519	ATTAATCTTGCAACTGCA	11537
RESULT 5			
US-09-949-016-5640			
; Sequence 5640, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5640			
; LENGTH: 2763			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-5640			
Query Match 81.2%; Score 515.8; DB 4; Length 2763;			
Best Local Similarity 91.2%; Pred. No. 8.1e-163;			
Matches 563; Conservative 0; Mismatches 42; Indels 12; Gaps 1;			
Qy	1	CCCTGTATCTTTAACTCTTGTAGTTTGTCTCTCCAGAACTCAAACTGTAAACTA	60
Db	2158	CCCTGTATCTTTAACTCTTGTATCTTGTCTCTCCAGAACTCAAACTGTAAACTA	2217
Qy	61	CAAAATTGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCC	120
Db	2218	-----CAAAATGGAGCCCAAGATGCAGTCCAGACTAAGATCTACCCGAGACCC	2265
Qy	121	CTGGACGGGCTCTAGCCCATGCTCGATGTTTAATGACATTTGAAGGCAACCCCTCCGAG	180
Db	2266	CTGGACGGGCTCTAGCCCATGCTGATGTTTAATGACATCAAGGCAACCCCTCCGAG	2325
Qy	181	GAATCTCACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACGGT	240
Db	2326	GAATCTCACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACGGT	2385
Qy	241	CATCAGCCAACTCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGACCTGAGAGAC	300
Db	2386	CGTCGGCCAACTCTCCCAACAGCACTTAGGTTTCTGTTGAGATGGGGACCTGAGAGAC	2445
Qy	301	AGACTAGCTGGATTTCTTAGCCCAACGAGAAATCCCTAAGCCTAGCTGGGAGGTGACT	360
Db	2446	AGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAGGTGACC	2505
Qy	361	GCATCCACCTCTAAACATGGGGTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTC	420
Db	2506	ACATCCACCTTTAAACAGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC	2565
Qy	421	ACTAAATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTTG	480
Db	2566	ACTAAATGCTAATTAGGCAAAAGACAGGAGGTAAAGAAATAGCAATCATCTATTGCTTG	2625
Qy	481	AGACACAGCGGAGGACAGAGTCCGGATATAACCCAGGCAATTCGAGCCGCAACGG	540
Db	2626	AGACACAGCAGGAGGACAGAAATGATCGGGATATAACCCAAAGTCTTCGAGCCGCAACGG	2685
Qy	541	CAACCCCTTTGGGTCCCTCTCTTTGTATGGGCGCTCTGTTTCACTATTTCACCTCT	600

Db	2686	CAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTTCATGCTATTTCACCTCT	2745
Qy	601	ATTAATCTTGCAACTG	617
Db	2746	ATTAATCTTGCAACTG	2762
RESULT 6			
US-09-949-016-15858/c			
; Sequence 15858, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 15858			
; LENGTH: 145320			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-15858			
Query Match 80.4%; Score 510.6; DB 4; Length 145320;			
Best Local Similarity 90.0%; Pred. No. 6e-160;			
Matches 575; Conservative 0; Mismatches 49; Indels 15; Gaps 2;			
Qy	1	CCCTGTATCTTTAACTCTTGTAAAGTTTGTCTCTTCCAGAATCAAACTGTAAACTA	60
Db	92601	CCCTGTATCTTTAACTCTTGTAAAGTTTGTCTCTTCCAGAATCAAAAGTTGTAAAGCTA	92542
Qy	61	CAAAATTGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC	120
Db	92541	CAAAATCGTTCTTCAAAATGGAAACCCAGATGAAGTCCATGACTAAGATCTACCGTGGACCC	92482
Qy	121	CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCACCCCTCCGAG	180
Db	92481	CTGGACCGGCTCTAGCCCATGCTCCCAATTTGTAATGATATCGAACGCAACCCCTCCGAG	92422
Qy	181	GAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACGGT	240
Db	92421	GAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACTGGT	92362
Qy	241	CATCAGCCAACTCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGACTGAGAGAC	300
Db	92361	CGTCAGCCAACTCTCCCAACAGCACTTGGGTTTTCTGTTGAGTGGGGGGACTGAGAGAC	92302
Qy	301	AGGACTAGCTGGATTTCTTAGGCCAAACGAAGAATCCCTAAGCCTAGCTGGGAAAGGTGACT	360
Db	92301	AGGATTAGCTGGATTTCTTAGGCCACTAAGAAATCCCAAGCCTAGCTGGGAAAGGTGACC	92242
Qy	361	GCATCCACCTCTAAACATGGGGTTGCAACTTAGCTCACCCCGACCAATC-----	411
Db	92241	ACATCCACCTTTAAACACTGGGCTTGCACCTTAGCTCACCCCGACCAATCAGGTAGTAA	92182
Qy	412	AGAGAGCTCACTAAATGCTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATC	471
Db	92181	AGAGAGCTCACTAAATGCTAATTTAGCAAAAAGAGGAGGTAAAAAATAGCCAATCATC	92122
Qy	472	TATTCGCTGAGAGCACAGCGGAGGGAACAAGGATCGGGATATAAACCCAGGATTCGAGC	531
Db	92121	TATCGCTGAGAGCACAGCGGGAAGGAACAATGATCGGGATATAAACCCAGGATTCGAGC	92062

QY 532 CGGCAACGGCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACTCTA 591
 Db 92061 CGGCAACGGCTACCTCTTTGGGTCCCTCTTGTATGGGGCTCTCTCT-----GT 92008
 QY 592 TTTCACTCTATTAAATCTTCAACTGAAAAA 630
 Db 92007 CTTCACTCTATTAAATCTTCAACTGAAAAAATA 91969

RESULT 7
 US-09-949-016-17411/c
 ; Sequence 17411, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17411
 ; LENGTH: 99580
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-17411

Query Match 71.3%; Score 453; DB 4; Length 99580;
 Best Local Similarity 88.9%; Pred. No. 1.2e-140;
 Matches 538; Conservative 0; Mismatches 55; Indels 12; Gaps 4;

QY 1 CCCTGTATCTTTAACTCTTTTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 60
 Db 6780 CCCTGTATCTTTAACTCTTTTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 6721
 QY 61 CAAATTTGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 6720 CAAATTTGTTCTTCAAAATAGAACCCAGATGAGTCCATGACTAAGATCCACCGTGGACCC 6661
 QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATGGAAGGACCCCTCCCGAG 180
 Db 6660 CTGGAAACAGCCTGCTAGCCCATGCTCCGATGTTAATGACATGGAAGGACCTCTCTCTGAT 6601
 QY 181 GAAATCTCACTGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTT-AGAGCGG 239
 Db 6600 GAAATCTCAACTGACACGACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAAGAGCGAG 6541
 QY 240 TCATCAGCAACCTCCCAACAGACACTTGGGTTTTCTTGTGAGAGGGGAGCTGAGAGA 299
 Db 6540 TCGTCGGCAACCTCCCAACAGACACTTGGGTTTTCTTGTGAGAGGGGAGCTGAGAGA 6481
 QY 300 CAGGACTAGCTGGATTTCTTAGGCAACAGAAATCCCTAAGCCCTAGCTGGG-AAAGTGA 358
 Db 6480 CAGGACTAGCTGGATTTCTTAGGCAACAGAAATTCCTAAGCCCTAGCTGGGAAAGTGA 6421
 QY 359 CTGATTCACCTCTTAAACATGGGGTTTCACTTACTAGCTACACCCGACCAATC----- 411
 Db 6420 CCGCACTACCTTTTAAACATGGGGTTTGTAACTCAGCTACACCCGACCAATCAGGTAGT 6361
 QY 412 --AGAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAG-CCAATC 468
 Db 6360 AAGAGAGGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGTCAATC 6301
 QY 469 ATCTATTCCCTGAGAGCACAGCGGAGGACAGGATCGGGATATAAACCAGGCAATTCG 528
 Db 6300 ATATATCGCTGAGAGCACAGGAGGAGCAATGATCAGGATATAAATCAGGCAATTCG 6241

QY 529 AGCCGCAACGGCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACT 588
 Db 6240 AGCAGGAGCAGCAACCCCTTTGGGTCCCTCTTGTATGGGGCTCTGTTTCACT 6181
 QY 589 CTATT 593
 Db 6180 CTATT 6176

RESULT 8
 US-09-949-016-13002/c
 ; Sequence 13002, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13002
 ; LENGTH: 168394
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(168394)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-13002

Query Match 70.6%; Score 448; DB 4; Length 168394;
 Best Local Similarity 88.7%; Pred. No. 8.4e-139;
 Matches 534; Conservative 0; Mismatches 55; Indels 13; Gaps 4;

QY 1 CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 60
 Db 25264 CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAATCAAACTGTAATACTA 25205
 QY 61 CAAATTTGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 25204 CAAATTTGTTCTTCAAAATGGAGCCCGACATGCAATGCAATCTTACTCGAGACCC 25145
 QY 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATGGAAGGACCCCTCCCGAG 180
 Db 25144 --GGACAGGCTGCTAGCCCATGCTCC-AGTTAATCAATCAAGGACCTCTCCAGAG 25088
 QY 181 GAAATCTCAACTGACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 Db 25087 GAAATCTCAACTGAAAGCCCATACTACACCCCAATTCAGCAGGAGCAGTTAGAGCAAT 25028
 QY 241 CATCAGCAACCTCCCAACAGACACTTGGGTTTTCTTGTGAGAGGGGAGCTGAGAGAC 300
 Db 25027 CATCGGCAATCT-CTCAACAGCACTTGGGTTTTCTTGTGAGAGGGGAGCTGAGAGAC 24969
 QY 301 AGGACTAGCTGGATTTCTTAGGCAAGAAATCCCTAAGCCCTAGCTGGGAAGGTGACT 360
 Db 24968 AGGACTAGCTGGATTTCTTAGGCGGCACTAAGAAATTCCTAAGCCCTAGCTGGGAAGGTGACC 24909
 QY 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 411
 Db 24908 GCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACATCTGACCAATCAGGTAGTAA 24849
 QY 412 AGAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATC 471

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCT-GGGAGGTGAC 359
 DB 15677 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAGGTGAC 15618
 QY 360 TGATCCACCTCTAAACATGGGCTTGCAACTAGCTCAGCCGACCAATC----- 411
 DB 15617 CACACCCACCTTTAAACATGGGCTTGCAACTAGCTCAGCCGACCAATCAGGTAGTA 15558
 QY 412 -AGGAGCTCAGCTAAATGCTTAATAGGC-AAAAATAGGAGGTAAAGAAATAGCCA-ATC 468
 DB 15557 AAGAGAGCTCAGCTAAATGCTTAATAGGC-AAAAATAGGAGGTAAAGAAATAGCATC 15498
 QY 469 ATCTATTGCTGAGAGCACAGCGGGAGGACAAAGGATCGGGATATAAACCCAGGCATTG 528
 DB 15497 ATCTGCTGCTGAGAGCACAGGGGAGGACAAATGATCAGATATAAACCCAGGCATTCA 15438
 QY 529 AGCGGCAACGGCAACCCCTTTGGGTCCCTCTCTTTGATGGGCTCTGTTTCACT 588
 DB 15437 AGCGGAATCGGCAACCCCTTTGGGTCCCTCTCTTTGATGGGAGCTCTGTTTCAAG 15378
 QY 589 CTATT 593
 DB 15377 ATATT 15373

RESULT 11
 US-09-949-016-44300/c
 ; Sequence 44300, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 44300
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-44300

Query Match 63.8%; Score 405.4; DB 4; Length 601;
 Best Local Similarity 88.1%; Pred. No. 5e-126;
 Matches 489; Conservative 1; Mismatches 52; Indels 13; Gaps 4;
 QY 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAAACTA 60
 DB 552 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAAACTA 493
 QY 61 CAAATTTGTTCTTAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
 DB 492 CAAATTTGTTCTTAAATGGAGCCCGAGATGCAGTCCATGACTAAGATCCACCGTGACCC 433
 QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGCACCCTCCCGAG 180
 DB 432 -GGACACGCTCTAGCCCATGCTCC-AGTTAATAATCAATCAAGGACCTCTCCAGAG 376
 QY 181 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 DB 375 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCAAT 316
 QY 241 CATCAGCAACCTCCCAACAGCACTTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC 300
 DB 315 CATCGGCCAATCT-CYCAACAGCACTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 257

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAGGTGACT 360
 DB 256 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAGGTGACC 197
 QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTAGCTCAGCCGACCAATC----- 411
 DB 196 GCATCCATCTTTAAACATGGGCTTGCAACTAGCTCAGCCGACCAATC----- 137
 QY 412 AGAGAGCTCAGCTAAATGCTTAATTAGGCCAAATAGGAGTAAAGAAATAGCCAATCATC 471
 DB 136 AGAGAGTTCAGCTAAATGCTTAATTAGGCCAAATAGGAGTAAAGAAATAGCCAATCATC 77
 QY 472 TATTGCTGAGAGCACAGCGGGAGGACAAAGGATCGGGATATAAACCCAGGCATTGAGC 531
 DB 76 TATTGCTGAGAGCACAGCGGGAGGAGTAATGATCAGGATATTAAGCCAGGCATTGAGC 17
 QY 532 CGGCAACGGCAACCC 546
 DB 16 TGGCAATGGCTACCC 2

RESULT 12
 US-09-949-016-13590/c
 ; Sequence 13590, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13590
 ; LENGTH: 149971
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)..(149971)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-13590

Query Match 57.8%; Score 367.2; DB 4; Length 149971;
 Best Local Similarity 80.9%; Pred. No. 1.4e-111;
 Matches 529; Conservative 0; Mismatches 83; Indels 42; Gaps 7;
 QY 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAAACTA 60
 DB 18750 CCCTGTATCTTTAACTCTTTGTTAAAGTTTCTCTTCCAGAAATCAAAACTGTAAAACTA 18691
 QY 61 CAAATTTGTTCTTAAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
 DB 18690 TAAATGTTCTTAAATGGAGCCCGCTGATGCA-----TTTGAATCTACTGCGGACCC 18640
 QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGCACCCTCCCGAG 180
 DB 18639 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGCACCCTCCCGAG 18580
 QY 181 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTT-AGAGCG 238
 DB 18579 GAAATCTCAATGACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAAGAGCG 18520
 QY 239 GTCATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAG 298

Db 18519 GTCGTGGCCAACTCCCAACAGCAGCTTGAGTTTCTGTTGAGCGGAGACTGAGAG 18460

Qy 299 ACAGGACTAGCTGGATTTCTTAGGCCAACAGAAATCCCTAGCTAGCTGGGAAGGTGA 358

Db 18459 ACAAGACTAGCTGGATTTCTTAGCGGACTAAGAATTCCTAAACCCAGCTGGGAAGGTGA 18400

Qy 359 CTGCATCCACTCTAAACATAGGGCTTGCAACTAGCTCACACCCGACCAATC-----411

Db 18399 CCGCACCCCACTTAAACACGGGGCTTGAACCTAGCTCACACCCCAACCAATCAGATAGT 18340

Qy 412 --AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGTAAAGAAATA-----G 462

Db 18339 AAGAGGGCTCACTAAATACAAATAGGCTAAGCAGGAGTAAAGAAAAAAGT 18280

Qy 463 CCAATCATCTATTGCTGAGAGCACAGCGGGAGGAGCAAGGATCGGGATATATAA-CCAG 521

Db 18279 CAAATCATATACCTGAGAGCATAGGGGGAGGACATGATGGGATATATAACCCCAA 18220

Qy 522 GCATTGAGCGGCAACGCCCTTTGGTCCCTCCCTTTGTTGATGGCGCTCTGT 581

Db 18219 GCATTTCTCGGGAGTGGAACCCCTTTGGTCTCTCCCTCCATGTTGATGGCTCTGTT 18160

Qy 582 TTTCACCTATTTCACCTATTAAATCTTGCAACTGMAAAAAAAGAAAAA 635

Db 18159 T-----CACTCTTAATCTTGCAATTGTGAGAAAAAACAATA 18120

RESULT 13

US-09-949-016-15497/c

; Sequence 15497, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15497

; LENGTH: 245286

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(245286)

; OTHER INFORMATION: n = A, T, C or G

US-09-949-016-15497

Query Match 46.0%; Score 292; DB 4; Length 245286;

Best Local Similarity 92.5%; Pred. No. 4.2e-86;

Matches 307; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 288 GGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAACGAAATCCCTAAGCCTAGC 347

Db 142457 GGTAGTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGCCTAGC 142398

Qy 348 TGGGAGGTGACTGCATCCACTCTAAACATGGGGCTTGCACCTTAGCTCACACCCGACC 407

Db 142397 TGGGAGGTGACTTCATCCACCTTTAAACATGGGGCTTGCACCTTAGCTCACACCCGACC 407

Qy 408 AATCAGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 467

Db 142337 AATCAGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 467

Qy 142337 AATCAGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 142278

Qy 468 CATCTATTGCTGAGAGCACAGCGGGAGGACAAGGATCGGGATATATAAACCCAGGCATTC 527

Db 142277 CATCTATCCCTGAGAGCACAGAGGGAGGACAATGATCAGGATATAAACCCAGGCATTC 142218

Qy 528 GAGCGGCAACGCAACCCCTTTGGTCCCTCCCTTTCTATGGGCGCTCTGTTTTCAC 587

Db 142217 CAGCCAGCAACGGCTACACTCTTTGGGTCCCTCCCTTTATATGGGAGCTCTGTTTTCAC 142158

Qy 588 TCTATTTTCACTCTATTAAATCTTGCAACTGAA 619

Db 142157 TCTATTTTCACTCTATTAAATCTTGCAACTGCA 142126

RESULT 14

US-09-949-016-133739/c

; Sequence 133739, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 133739

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-133739

Query Match 45.9%; Score 291.6; DB 4; Length 601;

Best Local Similarity 92.2%; Pred. No. 1.2e-87;

Matches 306; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

Qy 288 GGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAACGAAATCCCTAAGCCTAGC 347

Db 414 GGTAGTGAGAGACAGGACTAGCTGGATTTCTTAGCTGACTAAGAATCCCTAAGCCTAGC 355

Qy 348 TGGGAAGGTGACTGCATCCACTCTTAAACATGGGGCTTGCACCTTAGCTCACACCCGACC 407

Db 354 TGGGAGGTGACTTCATCCACTTTAAACATGGGGCTTGCACCTTAGCTCACACCCGACC 295

Qy 408 AATCAGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAT 467

Db 294 AATCAGAGAGCTCACTAAATGCTAAATAGGCAAAAGACAGGAGTTAAAGAAATAGCCAAT 235

Qy 468 CATCTATTGCTGAGAGCACAGCGGGAGGACAAGGATCGGGATATAAACCCAGGCATTC 527

Db 234 CATCTATCCCTGAGAGCACAGAGGGAGGACAATGATCAGGATATAAACCCAGGCATTC 175

Qy 528 GAGCGGCAACGCGCAACCCCTTTGGGTCCCTCCCTTTGTTGATGGGCGCTCTGTTTTCAC 587

Db 174 CAGCCAGCAACGGCTACACTCTTTGGGTCCCTCCCTTTATATGGGAGCTCTGTTTTCAC 115

Qy 588 TCTATTTTCACTCTATTAAATCTTGCAACTGAA 619

Db 114 TCTATTTTCACTCTATTAAATCTTGCAACTGCA 83

RESULT 15

US-09-949-016-15393

; Sequence 15393, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 15393
; LENGTH: 194790
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15393

Query Match 45.4%; Score 288; DB 4; Length 194790;
Best Local Similarity 82.6%; Pred. No. 8.1e-85;
Matches 393; Conservative 0; Mismatches 70; Indels 13; Gaps 5;
QY 130 CTTGCTAGCCCTCTCCGATGTTAATGACATTGAAGGACCCCTCCCGAGGAAATCTCA 189
DB 98345 CATAGTGGTCTTTTGGCAATTTCTGATGATCAATCAAGGCAATCCCTCTTGAGGAAATCTCA 98404
QY 190 ACTGCACAAACCCCTACTATGCCCAATTCAGCGGAAGCAATTAGAGCGGTCTATCAGCCA 249
DB 98405 ACTGCACAAACCCCTACTATGCCCAAGTTTACAGGAAAGCAATTTAGAGCAATTTGTGGCCA 98464
QY 250 ACCTCCCCCAACAGCAGCTTGGTTTCTGTTGAGAGGGGAGCTTGAGAGACAGGACTAGC 309
DB 98465 ACCTCCCCCAACAGCAGCTTGGTTTCTGTTGAGAGGGGAGCTTGAGAGACAGGACTAGC 98523
QY 310 TGGATTTCTTAGGCAACAGAAATCCCTAAGCTAGCT-GGGAAGGTGACTGCATCCAC 368
DB 98524 TGGATTTCTTAGGCAACAGAAATCCCTAAGCTAGCTGGGAAGGTGACTGCATCCAC 98583
QY 369 CTCTAAACATGGGCTTGCACCTTAGCTCAGCTCAGCCGACCAATC-----AGAGAGCT 419
DB 98584 CTTTAAACACAGAGCTTGTAACTCAGCTCAGCCGACCAATCAGGTAGTAAAGAGAGCT 98643
QY 420 CACTAAATATGTAATTAGGC-AAAAATAGGAGGTAAAGAAAT-AGCCAATCATCTATTGC 477
DB 98644 CACTAAATATGTAATTAGGC-AAAAATAGGAGGTAAAGAAAT-AGCCAATCATCTATCGC 98703
QY 478 CTGAGAGCAGACGCGGGAGGACAGGATCGGGATATAAACCCAGGCAATTCGAGCCGCAA 537
DB 98704 CTGAGAGCAGACGCGGGAGGACAGGATCGGGATATAAACCCAGGCAATTCGAGCCGATC 98763
QY 538 CGGCAACCCCTTTGGGTCCCTCCCTTTGATGGGCGCTCTGTTTCACTCTATT 593
DB 98764 AGGTAACCCCTTTGGGTCCCTCCCTCAGCTGTATGGGAGCTCTGTTTCACTCTATT 98819

RESULT 16
US-09-949-016-44296/c
; Sequence 44296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 44296
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44296
Query Match 43.0%; Score 273.2; DB 4; Length 601;
Best Local Similarity 85.5%; Pred. No. 1.9e-81;
Matches 376; Conservative 1; Mismatches 49; Indels 14; Gaps 6;
QY 167 GCACCCCTCCGAGGAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCAGCGGAA 226
DB 601 GCACCTCCTCTGAGGAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCAGCGGAA 542
QY 227 GCAGTT-AGAGGGGTCTATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTCTTTGAGAG 285
DB 541 GCAGTTAAGAGCGTGTTCGGCCAACTTCCCAACAGCAGCTTGGGTTTCTCTTTGAGAT 482
QY 286 GGGGAGCTGAGAGACAGGACTAGCTGGATTTCTAGGCCAACGAAATCCCTAAGCCTA 345
DB 481 GGGGAGCTGAGAGACAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATTCCTAAGCCTA 422
QY 346 GCT-GGGAAGGTGACTGCTATCCACTCTAAACATGGGCTTGCACCTAGCTCAGCCCG 404
DB 421 GCTGGGGAAGGTGACCGCACCCACCTTTAAACACAGGCGCTGGTAACTCAGCTCAGCCCTG 362
QY 405 ACCAATC-----AGAGAGCTCACTAAATGCTAAATTTAGGC-AAAAATAGGAGTAA 454
DB 361 ACCAATCGGGTAGTAAGAGGGCTCACTGAACACAAATTTAGGCTAAACAGGAGGATA 302
QY 455 AGAAATAG-CCAAATCATCTATTGCTGAGAGCAGCGGGAGGACAAAGATCGGATAT 513
DB 301 RGAACAGTCAATCATATATCGCTGAGAGCACAACAG-GGGACAATGATCGGATAT 243
QY 514 AAACCCAGGATTCGAGCGCGCAACCGCAACCCCTTTGGGTCCTCCCTTTGATGGG 573
DB 242 AAACCTCAGGCATTTCAAGCAGGAGTGGCAACCCCTTTGGGTCCTCCCTCGATGGG 183
QY 574 CGCTCTGTTTCACTCTATT 593
DB 182 AGCTCTGTTTCACTCTATT 163

RESULT 17
US-09-949-016-15086/c
; Sequence 15086, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 15086
; LENGTH: 219964
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15086

Query Match 42.4%; Score 269; DB 4; Length 219964;
Best Local Similarity 93.4%; Pred. No. 2.2e-78;
Matches 281; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 293 TGAGAGACAGGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTGGGA 352
Db 92554 TGAGAGACAGGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTGGGA 92495
QY 353 AGGTGACTGCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATCA 412
Db 92494 AGGTGACCACGTCACCTTTAAAGATGGGGCTTGCNACTTAGCTCACACCCGACCAATCA 92435
QY 413 GAGAGCTCACTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAAATAGCAATCATCT 472
Db 92434 GAGAGCTCACTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAAATAGCTAATCATCT 92375
QY 473 ATTGCTGAGACACAGCGGGAGGACAAAGATCGGGATATAAACCCAGGCAATCGAGCC 532
Db 92374 ATCGCTTAAGACACAGCGGGAGGACAAATGATCGGGATATAAACCCAGGCAATCGAGCC 92315
QY 533 GCGACGGCAACCCCTTTGGGTCCCTCTCCCTTTGATGGGCGCTCTGTTTTCACTCTAT 592
Db 92314 GCGACGGCTACCTCTTTGGGTCCCTCTCCCTTTGATGGGAGCTCTGTTCACTCTAT 92255
QY 593 T 593
Db 92254 T 92254
RESULT 18
US-09-573-080A-385
; Sequence 385, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 385
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(780)
; OTHER INFORMATION: ltr17
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-385
Query Match 41.4%; Score 262.8; DB 4; Length 780;
Best Local Similarity 93.2%; Pred. No. 7.2e-78;
Matches 289; Conservative 0; Mismatches 12; Indels 9; Gaps 1;
QY 293 TGAGAGACAGGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTGGGA 352
Db 1 TGAGAGACAGGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGCTGGGA 60
QY 353 AGGTGACTGCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATC- 411
Db 61 AGGTGACCGCTTCCACCTTTAAACACGGGGCTTGCNACTTAGCTCACACCCGACCAATCA 120
QY 412 -----AGAGAGCTCACTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAAATAGC 463

Db 121 GGTAGTAAAGAGAGCTCACTAAATGCTTAATTAGGCCAAAACAGGAGGTAAAGAAATAGC 180
QY 464 CAATCATCTATTTCCTTGAGAGCACAGCGGGAGGAGCAAGATCGGATATATAACCCAGGC 523
Db 181 CAATCATCTATTTCCTTGAGAGCACAGCGGGAGGAGCAATGATCGGATATATAACCCAGGC 240
QY 524 ATTTCGAGCGGCAACGCAACCCCTTTGGGTCCCTCTCCCTTTGATGGGCGCTCTGTTT 583
Db 241 ATTTCGAGCGGCAACGCTACCTCTTTGGGTCCCTCTCCCTTTGATGGGAGCTCTGTTT 300
QY 584 TCACTCTATT 593
Db 301 TCACTCTATT 310
RESULT 19
US-09-949-016-103638/c
; Sequence 103638, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103638
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-103638
Query Match 39.7%; Score 252.2; DB 4; Length 601;
Best Local Similarity 86.3%; Pred. No. 2.3e-74;
Matches 308; Conservative 1; Mismatches 29; Indels 19; Gaps 2;
QY 288 GGGACTGAGACACAGGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGC 347
Db 472 GGTGATGAGACACAGGACTAGCTGGATTCTTAGGCCAACAGAAATCCCTAGCCTAGC 413
QY 348 TGGGAAGGTGACTGATCCACTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACC 407
Db 412 TGGGAAGGTGACCACTCCACTCTTAAACACGCGGGCTTGTAACTTAGCTTACCCCGACC 353
QY 408 TATC-----AGAGAGCTCACTTAAATGCTTAATTAGGCCAAAATAGGAGGTAAAGAA 458
Db 352 AATCAGGTAGTAAAGAGAGCTCAATTAATGCTTAATTAGGGAACAGGARGTAAAGAA 293
QY 459 ATAGCAATCATCTATTTCCTTGAGAGCACAGCGGGAGGAGCAAGATCGGATATATAACC 518
Db 292 GTAGCCAATCATCTCTGCTGCTGAGAGCACAGGGGGAGGAGCAATGATCAGGATATATAACC 233
QY 519 GAGGATTTGAGCGCGGCAACCGCAACCCCTTTGGGTCCCTCTCCCTTTGATGGGCGCTC 578
Db 232 CAGTCATTCAGCTGGCAACCGCTTACCACCTTTGGGTCTCTCCCTTTGATGGGAGCTC 173
QY 579 TGTTTTCACTCTATTTCCTTAAATCTTTGCAACTTGTGCAACTGAAAAAATAAAAAA 635
Db 172 TGT-----TTTCACTCTATTAAATCTTTGCAACTGCAAAAAAATAAAAAA 126
RESULT 20
US-09-949-016-103639/c
; Sequence 103639, Application US/09949016

QY 408 AATC-----AGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAA 458
|||
Db 116297 AATCAGGTAGTAAGAGAGCTCATTAATAATGCTAATTAGGAAAACAGGAAGTAAAGAA 116238
|||
QY 459 ATAGCAATCATCTATTGCTCGAGAGCACAGCGGAGGACAAAGGATCGGATATAAAACC 518
|||
Db 116237 GTAGCCATCATCTGCTCGAGAGCACAGCGGAGGACAAATGATCAGGATATAAAACC 116178
|||
QY 519 CAGGCAATCGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGCGGCTC 578
|||
Db 116177 CAGTCATTCAGTGGCAACGCTACCCACTTTGGGTCTCTCTCCCTTTGTTATGGGAGCTC 116118
|||
QY 579 TCTTTTCACTCTATTTCACCTCTATTAAATCTTGCAACTGAAAAAAGAAAAA 635
|||
Db 116117 TGT-----TTTCACTCTATTAAATCTTGCAACTGCAAAAAAAGAAAAA 116071
|||

RESULT 23

US-09-949-016-44299/c
; Sequence 44299, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44299
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44299

Query Match 39.3%; Score 249.4; DB 4; Length 601;
Best Local Similarity 88.5%; Pred. No. 2e-73;
Matches 284; Conservative 1; Mismatches 27; Indels 9; Gaps 1;

QY 282 AGAGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAACGAAAGATCCCTAAG 341
|||
Db 601 AGAAGGAGACTGAGAGACAGGACTGACTGGATTTCTTAGGCCGACTAAGAATTCCTAAG 542
|||
QY 342 CCTAGCTGGGAAGTGACTGCATCCACCTTAACATGGGGCTTGCAACTTAGCTCACAC 401
|||
Db 541 CCTAGCTGGGAAGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAACATCAT 482
|||
QY 402 CCGACCAATC-----AGAGAGCTCACTAAATGCTAAATTAGCAAAAATAGGAGGT 452
|||
Db 481 CTGACCAATCAGGTAGTAAGAGAGTTCACTAAATGCTAACTAGCAAAAACAGAGGT 422
|||
QY 453 AAGAAATAGCCAAATCATCTATTGCTTGAGAGCACAGCGGAGGGAACAAGGATCGGGATA 512
|||
Db 421 AAGAAATAGCCAAATCATCTATTGCTTGAGAGCACAGCGGAGGGAATAATGATCAGGATA 362
|||
QY 513 TAAACCCAGGATTCAGCGCGCAACGCCCTTTGGGTCCCTCCCTTTGTTATGG 572
|||
Db 361 TAAACCCAGGATTTGAGCTGGCAATGGCTACCCCTCTTTGGGTCCCTCCCTTTGTTATGG 302
|||
QY 573 GCCTCTGTTTTCACCTCTATT 593
|||
Db 301 RAGCTCTGTTTTCACCTCTATT 281
|||

RESULT 24

US-09-949-016-17420

; Sequence 17420, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17420
; LENGTH: 84571
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84571)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17420

Query Match 38.7%; Score 245.8; DB 4; Length 84571;
Best Local Similarity 89.4%; Pred. No. 8e-71; 22; Indels 14; Gaps 3;
Matches 303; Conservative 0; Mismatches 0;

QY 291 ACTGAGACAGGACTAGCTGGATTTCTTAGGCCAACGAA-GAATCCCTAAGCCTTAGCTG 349
|||
Db 7322 ACTGAGACAGGACTAGCTGGATTTCTTAGGCCGACTAAGGAATCCCAAGCCTTAGCTG 7381
|||
QY 350 GGAAGGTGACTGCATCCACCTTAACATGGGGCTTGCAACTTAGCTCACACCCGACCAA 409
|||
Db 7382 GGAAGTGCACCGCTTCCACCTTTAAACATGGGACTTGCAACTTAGCTCACACCCGACCAA 7441
|||
QY 410 TCA-----GAGAGCTCACTAAATGCTAAATTAGGCAAAAATAGGAGGTAAAGAAAT 460
|||
Db 7442 TCAGATAGTAAGGAGAGCTCACTAAATGCTAAATTAGGCAAAAACAGGAGGTAAAGAAAT 7501
|||
QY 461 AGCCAAATCATCTATTGCTTGAGAGCACAGCGGAGGGAACAAGGATCGGGATATAAACCCA 520
|||
Db 7502 AGC-----CATCTGCTGCTTGAGAGCACAGGAGGAGCAATGATCAGGATATAAACCCA 7557
|||
QY 521 GGCATTCAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGCGCTCTG 580
|||
Db 7558 GGCATTCAGCGCAGCAATGCGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGCGGAGCTCTG 7617
|||
QY 581 TTTTCACTCTATTTCACCTCTAATAATCTTGCAACTGAA 619
|||
Db 7618 TTTTCACTGTGTTTCACTCTAATAATCTTGCAACTGCA 7656
|||

RESULT 25

US-09-949-016-17002
; Sequence 17002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 17068
 LENGTH: 140224
 TYPE: DNA
 ORGANISM: Human
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (1)...(140224)
 OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17068

Query Match 38.7%; Score 245.6; DB 4; Length 140224;
 Best Local Similarity 86.9%; Pred. No. 1.3e-70;
 Matches 285; Conservative 0; Mismatches 34; Indels 9; Gaps 1;
 QY 279 TTGAGAGGGGGAGCTGAGAGACGAGGACTAGCTGGATTTCCTAGGCCAACGAAGATCCCT 338
 Db 16074 TAGCAGGGCCCTATTTCAGAGACAGGACTAGCTGGATTTCCTAGGCCGACTAAGAATCCCT 16133
 QY 339 AAGCTAGCTGGGAGGTGACTGCATCCACTCTTAACATGCGGCTTGCACCTTAGCTCA 398
 Db 16134 AAGCTAGCTGGGAGGTGACTGCATCCACTCTTAACATGCGGCTTGCACCTTAGCTCA 16193
 QY 399 CACCGACCAATC-----AGAGAGCTCACTTAAATGCTAATAGGCAAAATAGGA 449
 Db 16194 CACCGACCAATCAGGTAGTAAGAGAGCTCACTTAAATGCTAATAGGCAAAATAGGA 16253
 QY 450 GGTAAGAAATAGCAATCATCTATTGCTGAGAGCACAGCGGAGGACAAGATCGGG 509
 Db 16254 GGTAAGAAATAGCAATCATCTATTGCTGAGAGCACAGCGGAGGACAAGATCGGG 16313
 QY 510 ATATAACCCAGGATTCGAGCGGCAACGCGCAACCCCTTTGGTCCCTCCCTTTGTA 569
 Db 16314 ATATAACCCAGGATTCGAGCGGCAACGCGCTACCCCTTTGGTCCCTCCCTTTGTA 16373
 QY 570 TGGCGCTCTGTTTCACTCTATTTCAC 597
 Db 16374 TGGAGCTCCGTTTCACTCTATTTCAC 16401

RESULT 26
 US-09-949-016-17068/c
 ; Sequence 17068, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17068
 ; LENGTH: 89584
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-17068

Query Match 37.7%; Score 239.4; DB 4; Length 89584;
 Best Local Similarity 89.8%; Pred. No. 1.2e-68;
 Matches 283; Conservative 0; Mismatches 21; Indels 11; Gaps 2;
 QY 288 GGGAGCTGAGACAGGAGCTAGCTGGATTTCCTAGGCCAACGAAGATCCCTTAAGCCTAGC 347
 Db 66011 GGAATTAGACAGGAGCTAGCTGGATTTCCTAGGCCAACGAAGATCCCTTAAGCCTAGC 65952

QY 348 TGGGAAGGTGACTGCATCCACCTCTAAACATCGGGCTTGCACCTTAGCTCACACCCGACC 407
 Db 65951 TGGGAAGGTGATCACATCCACCTTTAAACATCGGGCTTGCACCTTAGCTCACACCCGACC 65892
 QY 408 AATCAG-----AGAGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAA 458
 Db 65891 AATCAGGTAGTAAGAGAGAGCTCACTAAATGCTAATAGGCAATAA--AGAGGTAAGAA 65834
 QY 459 ATAGCCAATCATCTATTGCTGAGAGCACAGCGGAGGAGCAAGGATCGGGATATAAAC 518
 Db 65833 ATAGCCAATCATCTATTGCTGAGAGCACAGCGGAGGAGCAATGATCGGGATATAAAC 65774
 QY 519 CAGGCAATTCAGCGCGCAACCGCAACCCCTTTGGGTCCCTCCCTTTGATGGGCGCTC 578
 Db 65773 CAGGCAATTCAGCGCGCAATGCTACCCCTTTGGGTCCCTCTCTTTGATGGGAGCTC 65714
 QY 579 TGTTTCACTCTATT 593
 Db 65713 TGTTTCACTCTATT 65699

RESULT 27
 US-08-686-878A-50
 ; Sequence 50, Application US/08686878A
 ; Patent No. 5708157
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John
 ; APPLICANT: Lavallie, Edward
 ; APPLICANT: Racie, Lisa
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Spaulding, Vikki
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 ; NUMBER OF SEQUENCES: 71
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/686,878A
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brown, Scott A.
 ; REGISTRATION NUMBER: 32,724
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8224
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 50:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 279 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-686-878A-50

Query Match 36.8%; Score 233.6; DB 1; Length 279;
 Best Local Similarity 88.1%; Pred. No. 2.6e-68;
 Matches 245; Conservative 1; Mismatches 32; Indels 0; Gaps 0;


```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 128750
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-128750

Query Match
Best Local Similarity 30.8%; Score 195.6; DB 4; Length 601;
Matches 280; Conservative 1; Mismatches 55; Indels 12; Gaps 4;

QY 130 CTGTCTAGCCCTGCTCCGATGTTAATGACATTGAAGGCACCCCTCCGAGGAAATCTCA 189
DB 242 CATAGTGGTCCTTTTGCCATTTCTGATGACATCAAGGCATCCCTCTTGAGGAAATCTC 301

QY 190 ACTGCACAAACCCCTACTATGCCCCAATTCAGCGGGAAGCAGTAGTTAGACGGTTCATCAGCCA 249
DB 302 ACTGCACAAACCCCTACTATGCCCCCAGTTTCAGCAGGAAGCAGTAGTTAGACGAGTTGTTGGCCA 361

QY 250 ACTCCCCCAACAGCACTTCGGTTTTCCTGTTGAGGGGGGACTGCAGACAGGACTAGC 309
DB 362 ACCTCCCAACAGCACTTCGGTTTTCCTGTTGAGGGGGGACTGCAGACAGGACTAGC 420

QY 310 TGGATTTCCTAGGCCAACGAAGAATCCCTAAGCTTAGCT-GGGAAGGTGACTGCATCCAC 368
DB 421 TGGATTTCCTAGGCCAACGAAGAATCCCTAAGCTTAGCT-GGGAAGGTGACTGCATCCAC 480

QY 369 CTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC-----AGAGAGCT 419
DB 481 CTTTAAACACAGAGCTTGTAACTCAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCT 540

QY 420 CACTAAATGCTTAATTAGGC-AAAAATAGGAGGTAAAGAAATAGCCAA 466
DB 541 CACTAAATACCAATTAGGCTAAAGAAACAGGAGGTAAAGAAATATCA 598

RESULT 32
US-09-949-016-17057
; Sequence 17057, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17057
; LENGTH: 154023
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-17057

Query Match
Best Local Similarity 28.4%; Score 180.4; DB 4; Length 154023;
Matches 271; Conservative 0; Mismatches 46; Indels 13; Gaps 5;

QY 293 TGAGAGACAGGCTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCT-GGG 351
DB 92234 TGAGAGACAGGCTAGCTGGATTTCCTAGGGCTGACTAAGAATTCCTAAGTCTATCTGGG 92293

QY 352 AAGTGACTGCATCCACTCTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC 411
DB 92294 AAGTGACTGGGCCCACTTTAAACATGGGGCTTGTAACCTCAGCTCACACCTGACCAATC 92353

```



```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;
; FILE OF INVENTION: CLO01307
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; CURRENT FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; CURRENT FILING DATE: 2000-09-08
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195792
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-195792

```

Query Match	27.8%;	Score 176.8;	DB 4;	Length 601;	
Best Local Similarity	85.4%;	Prod. No. 6e-49;	32;	Indels 14;	Gaps 6;
Matches 270;	Conservative 0;	Mismatches 270;			
Qy	292	CTGAGAGCAGGACTAGCTGGATTTCCTTAGGCCAACGAAGAATCCCTAAGCCTAGCT-GG	350		
Db	540	CTGAGAGCAGGACTAGCTGGATTTCCTTAGGCTGACTAGGAATCTTAAAGCCTAGCTGGG	481		
Qy	351	GAAGGTGACTGCATCCACCTCTAAACATGGGGCTTGGCAACTTAGCTCACACCCGACCAAT	410		
Db	480	GAAGGTGACCGCACCCACCTTTAAACACGGAGCTTGTAACCTCAGCTCACACCCACCAAT	421		
Qy	411	C-----AGAGAGCTCACTAAATCTCTAATTAGGC-AAAAATAGGAGGTAAAGAAAT	460		
Db	420	CAGGTAGTAAAGAGGGCTCCTAAATAACAAATTAGGCTTAAGACGAGGAGGTAAAGAAAT	361		
Qy	461	AG-CCAATCATCTATTGCTCGAGAGCACAGCGGGAGGGACAAAGGATCGGGATATAAAC-	518		
Db	360	AGTCAAATCATCTATCTCGAGAGCACAGGGGGAGGGNCATGATTGGGATATAAACCCY	301		
Qy	519	CAGGCATTTCGAGCCGGCA-ACGGCAACCCCTTTTGGGTCCCTCCCTTTGTATGGGCGCT	577		
Db	300	CAGGCATTTCGAGCCGGGATGGGCAACCCCTTTTGGGTCCCTCCCTTTGTATGGGAGCT	241		
Qy	578	CTGTTTTCACTCTATT	593		
Db	240	CTGTTTTCACTCTATT	225		

```

RESULT 35
US-09-949-016-17296
; Sequence 17296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH POLYOMYELITIS WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17296
; LENGTH: 251672
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature

```

Qy	412	-----RGAGAGCTCACTAAATGCTTAATTAGGC-AAAAATAGGAGGTAAAGAAATA	461
Db	92354	AGGTAATAAGAGAAGTCATTAAATACCAATTAGGCTAAAGCAGGAGTAAAGAAATA	92413
Qy	462	GCCAA-TCATCTATTGGCTGAGAGCACA-GCGGAGGAGCAAGGATCGGGATATAAAACC	519
Db	92414	GTCAAGTCATCTATCGCTGAGAGTACACTGGGAGGAGCAATGATCGGATATAAACTC	92473
Qy	520	AGGCATTGAGCGCGCAACGCGCAACCCCTTTGGGTCCCTTCCCTTTGTATGGCGGCTCT	579
Db	92474	AGGCATTGAAATCGGCAGTGCAAAACCCCTTTGAGTCOCCTCCCGTTGTATAGGAGCTCT	92533
Qy	580	GTTTTCACCTCTATTTCACCTCTATTAAATCT	609
Db	92534	GTTTTCACTCTGTTAAATAAATCTTGCACCT	92563

```

RESULT 33
US-09-949-016-44301/c
; Sequence 44301, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44301
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44301

```

Query Match	28.0%;	Score 178;	DB 4;	Length 601;
Best Local Similarity	88.5%;	Prod. No. 2.4e-49;		
Matches 216;	Conservative 0;	Mismatches 25;	Indels 3;	Gaps 2;
Qy 1	CCCTGTATCTTTAAACCTCCTTGTTAAGTTTGTCCTTCCAGAATCAAAACTGTFAAAACTA	60		
Db 242	CCCTGTATCTTTAAACATCCTTGTTAAGTTTGTCCTTCCAGAATCAAAAGCTATAAAACTA	183		
Qy 61	CAAAATGTTCTTCAATGAGCACCAGATGGATGCCATGACTAAGATCCACCGTGACCC	120		
Db 182	CAAAATGTTCTTCAATGAGGCCAGATGCGATGCCATGACTAATCTACTGCAGACCC	123		
Qy 121	CTGGACCGCGCTGCTAGCCCATGCTCCGATGTTTAATGACATTGAAGGACCCCTCCCGAG	180		
Db 122	--GGACCGCTGCTAGCCCATGCTCC--AAGTTTAATGAATCATCAAGGACACCTCTCCAGAG	66		
Qy 181	GAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGGAAGACAGTTAGAGCGGT	240		
Db 65	GAATCTCAACTGACAGCCCATCTACACCCCAATTCAGCAGGAGCAGTTAGAGCAAT	6		
Qy 241	CATC 244			
Db 5	CATC 2			

RESULT 34
US-09-949-016-195792/c
; Sequence 195792, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.


```
US-09-949-016-13413/c
; Sequence 13413, Application US/09949016
; Patent No. 6812339
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13413
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13633
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13633
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13633
Query Match 27.4%; Score 174; DB 4; Length 39686;
Best Local Similarity 83.1%; Pred. No. 7.9e-47;
Matches 261; Conservative 0; Mismatches 40; Indels 13; Gaps 5;
QY 293 TGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAAGAAATCCCTAAGCCTAGCT-GGG 351
Db 6761 TGAGAGGCGAGGACTAGCTGGATTTCCTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGGG 6702
QY 352 AAGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC 411
Db 6701 AAGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCTCAGCTCCGACCAATC 6642
QY 412 -----AGAGAGCTCAGCTAAATGCTAATAGG-CAAAAAATAGGAGGTAAAGAAATA 461
Db 6641 AGTGTTGTAAGAGGCTCAGCTAAATGCTAATAGGTTAAAGCAGGAGGTAAAGAAATA 6582
QY 462 -GCCAATCATCTATTGCTGAGAGCAGCGGGAGGAGCAAGGATCGGATATATAA-CCC 519
Db 6581 CTCAAATCATCTATCATCTGAGAGCAGCGGGAGGAGCAAGGATCGGATATATAA-CCCC 6522
QY 520 AGGCAATTCGAGCGGCAACGCCACCCCTTTGGTCCCTCCCTTTGTATGGCGCTCT 579
Db 6521 AGGCATTTGAGCAAGGAGGGGCAACCCCTTTGGATCCCTACCTTGTATAGCAGCTCT 6462
QY 580 GTTTTCACTCTATT 593
Db 6461 GTTTTCACTCTATT 6448
RESULT 40
US-09-949-016-11770/c
; Sequence 11770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11770
; LENGTH: 49487
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(49487)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11770
Query Match 27.4%; Score 174; DB 4; Length 49487;
Best Local Similarity 83.1%; Pred. No. 9.1e-47;
US-09-949-016-13413/c
; Sequence 13413, Application US/09949016
; Patent No. 6812339
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13413
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13633
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13633
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13633
Query Match 27.5%; Score 174.6; DB 4; Length 116652;
Best Local Similarity 84.9%; Pred. No. 1e-46;
Matches 269; Conservative 0; Mismatches 34; Indels 14; Gaps 6;
QY 291 ACTGAGAGCAGGACTAGCTGGATTTCCTAGGCCAACGAAGAAATCCCTAAGCCTAGCT-G 349
Db 6704 AGTGAGAGCAGGACTAGCTGGATTTCCTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGG 66981
QY 350 GGAAGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAA 409
Db 66980 GGAAGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCTGACCAA 66921
QY 410 TC-----AGAGAGCTCAGCTAAATGCTAATAGG-CAAAAAATAGGAGGTAAAGAAA 459
Db 66920 TCAGGTAGTAAAGGGCTCAGCTAAATGCTAATAGGTTAAAGCAGGAGGTAAAGAAA 66861
QY 460 TAG-CCAATCATCTATTGCTGAGAGCAGCGGGAGGAGCAAGGATCGGATATATAA-C 517
Db 66860 TAGTCAATAATCTATCATCTGAGAGCAGCGGGAGGAGCAAGGATCGGATATATAACC 66801
QY 518 CCAGGCATTCGAGCGGCA-ACGCCACCCCTTTGGTCCCTCCCTTTGTATGGCGC 576
Db 66800 CCAGGCATTCGAGCTGGAGTGGGCAACCCCTTTGGTCCCTCCCTTCGATGGGAGC 66741
QY 577 TCTGTTTCACTCTATT 593
Db 66740 TCTGTTTCACTCTATT 66724
RESULT 39
US-09-949-016-13633/c
; Sequence 13633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

	Matches	261;	Conservative	0;	Mismatches	40;	Indels	13;	Gaps	5;
Qy	293	TGAGAGACAGGACTAGCTGGATTTCCTAGGCCAACGAGAAATCCCTAAGCCTAGCT-GGG	351							
Db	16051	TGAGAGGCGAGGACTAGCTGGATTTCCTAGGCCGACTAAGAAATTCCTAAGCCTAGCTGGGG	15992							
Qy	352	AAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTTAGCTCACACCCGACCAATC	411							
Db	15991	AAGGTGACTGCACCCACCTTTAAACACGGGGCTTGTAACCTCAGCTCACACTCGACCAATC	15932							
Qy	412	-----AGAGAGCTCACTAAATGCTTAATTAGG-CAAAATAGGAGGTAAGAAATA	461							
Db	15931	AGGTGGTAAAGAGGCTCACTAAATATCAATTAGGTTAAAGCAGGAGGTAAGAAATA	15872							
Qy	462	-GCCAATCATCTATTGCCTGAGAGCACAGCGGGGACAGGATCGGGATATAAA-CCC	519							
Db	15871	CTCAATCATCTATCTATCTGAGAGCACAGGGGGAGCGCAATGATTGGGATATAAACCCC	15812							
Qy	520	AGGCATTTCGAGCGGGCAACCGCAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCT	579							
Db	15811	AGGCATTTCGAGCGGGCAACCGCAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCT	15752							
Qy	580	GTTTTCACCTATT	593							
Db	15751	GTTTTCACCTATT	15738							

Search completed: February 21, 2005, 15:01:18
 Job time : 120.885 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 11:52:50 ; Search time 350.955 Seconds
(without alignments)
10694.149 Million cell updates/sec

Title: US-09-319-156B-6

Perfect score: 635

Sequence: 1 cccgtatcttttaacctctt.....tgaaaaaaaaaaaaaaaaaa 635

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10760316

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	635	100.0	635	8	US-08-979-847-102
2	635	100.0	635	16	US-10-114-104-102
3	617.4	97.2	2030	17	US-10-637-793-25
4	564.8	88.9	366710	18	US-10-719-993-6792
5	546.4	86.0	2074	17	US-10-416-642-4
6	545	85.8	1329	8	US-08-979-847-108
7	545	85.8	1329	16	US-10-114-104-108
8	543.4	85.6	21646	13	US-10-087-192-910
9	533.8	84.1	2946	13	US-10-114-893-134
10	533.8	84.1	2946	15	US-10-016-249-3
11	532.8	83.9	326014	9	US-09-731-231A-3

c	12	532.8	83.9	326014	18	US-10-751-985-3	Sequence 3, Appli
	13	525.8	82.8	1136	17	US-10-632-793-25	Sequence 25, Appli
	14	522.8	82.3	283351	18	US-10-719-993-7065	Sequence 7065, Ap
	15	522.8	82.3	283351	19	US-10-741-600-17995	Sequence 17995, A
	16	522.8	82.3	1980090	18	US-10-719-993-6815	Sequence 6815, Ap
	17	522.8	82.3	1980090	19	US-10-741-600-17676	Sequence 17676, A
	18	522.6	82.3	2782	17	US-10-632-793-26	Sequence 26, Appli
	19	521	82.0	2782	17	US-10-133-036-1	Sequence 1, Appli
	20	517.4	81.5	2930	9	US-09-802-535-1	Sequence 1, Appli
	21	516.2	81.3	56093	10	US-09-873-367C-81	Sequence 81, Appli
	22	511.4	80.5	7582	17	US-10-632-793-30	Sequence 30, Appli
	23	510.6	80.4	161334	13	US-10-087-192-730	Sequence 730, App
c	24	500.8	78.9	285020	13	US-10-087-192-1666	Sequence 1666, Ap
	25	499.6	78.7	1894	9	US-09-864-761-4444	Sequence 4444, Ap
	26	495.6	78.0	3372	17	US-10-632-793-28	Sequence 28, Appli
	27	493.2	77.7	2052	17	US-10-276-774-678	Sequence 678, App
	28	482.6	76.0	22436	11	US-09-997-723-148	Sequence 148, App
	29	481.4	75.8	849	17	US-10-220-120-15	Sequence 15, Appli
	30	466	73.4	2349	17	US-10-104-047-1148	Sequence 1148, Ap
	31	449	70.7	1564	15	US-10-204-887-10	Sequence 10, Appli
	32	446.2	70.3	1393	17	US-10-220-120-17	Sequence 17, Appli
	33	434.4	68.4	7974	14	US-10-198-846-9936	Sequence 9936, Ap
	34	427.6	67.3	44063	18	US-10-322-281-718	Sequence 718, App
	35	424.6	66.9	604	13	US-10-027-632-289680	Sequence 289680,
c	36	424.6	66.9	604	17	US-10-027-632-289680	Sequence 289680,
	37	420.4	66.2	50353	19	US-10-741-600-17699	Sequence 17699, A
	38	415.6	65.4	1394	18	US-10-466-531-47	Sequence 47, Appli
	39	395.8	62.3	548	16	US-10-029-386-8554	Sequence 8554, Ap
c	40	388.4	61.2	23855	17	US-10-242-355-1175	Sequence 1175, Ap
	41	383.4	60.4	657	13	US-10-027-632-224966	Sequence 224966,
c	42	383.4	60.4	657	17	US-10-027-632-224966	Sequence 224966,
c	43	373	58.7	439	9	US-09-764-881-183	Sequence 183, App
c	44	373	58.7	439	10	US-09-764-881-183	Sequence 183, App
c	45	373	58.7	439	17	US-10-242-747-183	Sequence 183, App

ALIGNMENTS

RESULT 1
US-08-979-847-102
; Sequence 102, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUBE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435


```

QY 361 GCATCCACCTCTAAACATGGGCTTGCACCTTAGCTCACACCGACCAATCAGAGAGCTC 420
DB 361 GCATCCACCTCTAAACATGGGCTTGCACCTTAGCTCACACCGACCAATCAGAGAGCTC 420
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCTG 480
DB 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCCTG 480
QY 481 AGAGCAGCGGGAGGGAAGAAGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 540
DB 481 AGAGCAGCGGGAGGGAAGAAGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 540
QY 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
QY 601 ATTAATCTTGCACCTGAAAAAAGAAAAA 635
DB 601 ATTAATCTTGCACCTGAAAAAAGAAAAA 635

RESULT 3
US-10-637-565-18
; Sequence 18, Application US/10637565
; Publication No. US20040043381A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: PERRON, Hervé
; APPLICANT: KOMURIAN-PRADEL, Florence
; TITLE OF INVENTION: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES
; TITLE OF INVENTION: METHODS FOR DETECTING MSRV-1 RETROVIRUS
; FILE REFERENCE: 110257
; CURRENT APPLICATION NUMBER: US/10/637,565
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US/09/890,340
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/IB00/00159
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: EP 99420041.8
; PRIOR FILING DATE: 1999-02-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 2030
; TYPE: DNA
; ORGANISM: MSRV-1 retrovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1626)
US-10-637-565-18

Query Match 97.2%; Score 617.4; DB 17; Length 2030;
Best Local Similarity 98.3%; Pred. No. 1.3e-181;
Matches 624; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTTGTGTTAGTTGTCTTTCCAGAAATCAAAAATGTAAGACTA 60
DB 1396 CCTGTATCTTTAACTCTTGTGTTAGTTGTCTTTCCAGAAATGTAAGACTA 1455

QY 61 CAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAGATCCACCGTGGACCC 120
DB 1456 CAGATGGTCTTACAAATGGAAACCCAGATGGAGTCCATGACTAGATCCACCGTGGACCC 1515

QY 121 CTGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGCACCCTCCCGAG 180
DB 1516 CTGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGCACCCTCCCGAG 1575

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 1576 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 1635

QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300

```

```

DB 1636 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGACTGAGAGAC 1695
QY 301 AGGACTAGCTGATTTCTTAGGCCAACGAAGAATCCCTAAGCTTAGCTGGGAAGGTGACT 360
DB 1696 AGGACTAGCTGATTTCTTAGGCCAACGAAGAATCCCTAAGCTTAGCTGGGAAGGTGACT 1755
QY 361 GCATCCACCTCTAAACATGGGCTTGCACCTTAGCTCACACCGACCAATCAGAGAGCTC 420
DB 1756 GCATCCACCTCTAAACATGGGCTTGCACCTTAGCTCACACCGACCAATCAGAGAGCTC 1815
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCCTG 480
DB 1816 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCCTG 1875
QY 481 AGAGCAGCGGGAGGGAAGAAGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 540
DB 1876 AGAGCAGCGGGAGGGAAGAAGATCGGATATAAACCCAGGATTCGAGCCGGCAACGG 1935
QY 541 CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
DB 1936 CAACCCCTTTGGGTCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 1995
QY 601 ATTAATCTTGCACCTGAAAAAAGAAAAA 635
DB 1996 ATTAATCTTGCACCTGAAAAAAGAAAAA 2030

RESULT 4
US-10-719-993-6792/c
; Sequence 6792, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6792
; LENGTH: 366710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(366710)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-719-993-6792

Query Match 88.9%; Score 564.8; DB 18; Length 366710;
Best Local Similarity 93.4%; Pred. No. 3.9e-164;
Matches 590; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 CCTGTATCTTTAACTCTTGTGTTAGTTGTCTTTCCAGAAATCAAAAATGTAAGACTA 60
DB 43899 CCTGTATCTTTAACTCTTGTGTTAGTTGTCTTTCCAGAAATGTAAGACTA 43840

QY 61 CAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAGATCCACCGTGGACCC 120
DB 43839 CAATGTCTTCAATGGAAACCCAGATGCAGTCCATGTCTAAGATCTCCACGGACCC 43780

QY 121 CTGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAGCACCCTCCCGAG 180
DB 43779 CTAGACTGGCTGCTAGCCCATGCTCTGGTGTGTTAATGACATCGAAGTCACTCTCTCTGAG 43720

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
DB 43719 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 43660

QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300

```


Db 43659 CATCAGCCAACTCACCACAGCACTTGGGTTTTCTGTGTGAGAGGGGACTGAGAGAC 43600
QY 301 AGGACTAGCTGATTTCTTAGGCCAAAGAGAAATCCCTAAGCTAGCTGGGAAGTGACT 360
Db 43599 AGGACTAGCTGATTTCTTAGGCCAAAGAGAAATCCCTAAGCTAGCTGGGAAGTGACT 43540
QY 361 GCATCCACTCTAACAATGGGGCTTGCACCTTAGCTTCAACCCGACCAATCAGAGAGCTC 420
Db 43539 GCATCCACTCTAACAATGGGGCTTGCACCTTAGCTTCAACCCGACCAATCAGAGAGCTC 43480
QY 421 ACTAAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCTG 480
Db 43479 ACTAAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCTG 43420
QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGAGCTTCAAGCCGGCAACGG 540
Db 43419 AGAGCACAGTGGGAGGAGCAAGGATTCGAATATAAACCAGGAGCTTCAAGCCGGCAACGG 43360
QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 43359 CAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTCACTCTATTTCACCTCT 43300
QY 601 ATTAATCTTGCACTGAAAAAATGAAAAA 632
Db 43299 ATTAATCTTGCACTGAAAAAATGAAAAA 43268

RESULT 5
US-10-416-642-4
; Sequence 4, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: ARVIZU, Chandra
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0842 PCT
; CURRENT APPLICATION NUMBER: US/10/416,642
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/249,407
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7477736CB1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1995
; OTHER INFORMATION: a, t, c, g, or other
US-10-416-642-4
Query Match 86.0%; Score 546.4; DB 17; Length 2074;
Best Local Similarity 93.8%; Pred. No. 1.9e-159;
Matches 579; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
QY 1 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTTCAGAGATCAAAATCTGTAAACTA 60
Db 1459 CCTGTATCTTTAACTCTTGTAAAGTTTGTCTTTCAGAGATTCGAGCTGTAAAGCTA 1518
QY 61 CAAATTTGTTCTTCAAAATGGAGCACCAAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
Db 1519 CAAATGTTCTTCAAAATGGAGCCCAAGATGAGTCCATGACTTAAATCTTACCAGGACCC 1578
QY 121 CTGAGCGGCTCTAGCCCATGCTCGATGTTTAATGACATTAAGAGCCCTCCCGAG 180
Db 1579 CTGAGCGGCTCTAGCCCATGCTCGATGTTTAATGACATTCGAGGCACTCTCCCAAG 1638
QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240

Db 1639 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGAAGCAGTTAGAGCGGT 1698
QY 241 CATCAGCAACCTCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGGGACTGAGAGAC 300
Db 1699 CGTCAGTCAACCTCCCAACAGCACTTGGGTTTTCTGTGAGAGGGGGGACTGAGAGAC 1758
QY 301 AGGACTAGCTGATTTCTTAGGCCAAAGAGAAATCCCTAAGCTAGCTGGGAAGTGACT 360
Db 1759 AGGACTAGCTGATTTCTTAGGCCAAAGAGAAATCCCTAAGCTAGCTGGGAAGTGACT 1818
QY 361 GCATCCACTCTAACAATGGGGCTTGCACCTTAGCTTCAACCCGACCAATCAGAGAGCTC 420
Db 1819 CGTCCCACTTTAAACACGGGGCTTGCACCTTAGCTTCAACCCGACCAATCAGAGAGCTC 1878
QY 421 ACTAAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCTG 480
Db 1879 ACTAAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATCTATTGCGCTG 1938
QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 1998 CAACCCCTTTGGTCCCTCCCTTTGTATGGAGCTCTGTTTCACTCTATTTCACCTCT 2057
QY 601 ATTAATCTTGCAACTG 617
Db 2058 ATTAATCTTGCAACTG 2074

RESULT 6
US-08-979-847-108
; Sequence 108, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACALUA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787

Db 19931 GCATCCACCTTTAAACATAGGGCTTGCACCTAGCTCACACCAACCAATCAGAGAGCTC 19990
 Qy 421 ACTAAATGCTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG 480
 Db 19991 ACTAAATGCTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTG 20050
 Qy 481 AGAGCACAGCGGAGGAGCAAGGATCGGATATAAAACCCAGGCAATTCGAGCCGCAACGG 540
 Db 20051 AGAGCACAGCGGAGGAGCAAGGATCGGATATAAAACCCGCGCATTCAGCCGCAATGG 20110
 Qy 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTCTACTCTATT 593
 Db 20111 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTCTACTCTATT 20163

RESULT 9

US-10-114-893-134
 ; Sequence 134, Application US/10114893
 ; Publication No. US20020193567A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Bowman, Michael R.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Carlin-Duckett, McKeough
 ; APPLICANT: Kelleher, Kerry S.
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: GI 6000-10A
 ; CURRENT APPLICATION NUMBER: US/10/114,893
 ; CURRENT FILING DATE: 2002-04-02
 ; EARLIER APPLICATION NUMBER: 09/413,232
 ; EARLIER FILING DATE: 1999-10-06
 ; NUMBER OF SEQ ID NOS: 321
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 134
 ; LENGTH: 2946
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-114-893-134

Query Match 84.1%; Score 533.8; DB 13; Length 2946;
 Best Local Similarity 91.5%; Pred. No. 1.9e-155;
 Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;
 Qy 1 CCCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60
 Db 2323 CCCTGTATCTTTAACTCTCTTTGTTAACTTGTCTCTCCAGAACTCAAACTGTAAACTA 2382
 Qy 61 CAATTTGTTCTTCAATGGAGCACAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 120
 Db 2383 -----CAAATGGAGCCCAAGATGAGTCCAAAGATTAAGATCAAAAGGCACCCCTCTGAG 2430
 Qy 121 CTGAGCCGCTGTAGCCATGCTCCGATGTTAATGATTAAGAGGCAACCCCTCCCGAG 180
 Db 2431 CTGAGCCGCTGTAGCCATGCTCCGATGTTAATGATTAAGAGGCAACCCCTCTGAG 2490
 Qy 181 GAAATCTCAATGCAACCCCTCTATATGCCCAATTCAGCGGAGCAAGTTAGAGCGGT 240
 Db 2491 GAAATCTCAATGCAACCCCTCTATATGCCCAATTCAGCGGAGCAAGTTAGAGCGGT 2550
 Qy 241 CATCAGCAACCTCCCAACAGCACTTGGTGTTCCTGTTGAGAGGGGAGCTGAGAGAC 300
 Db 2551 CGTCGCGCAACCTCCCAACAGCACTTGGTGTTCCTGTTGAGATGGGAGCTGAGAGAC 2610
 Qy 301 AGGACTAGCTGGATTTCTTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 2611 AGGACTAGCTGGATTTCTTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 2670

Db 677 ACTAAATGCTAATCAGGCAAAACAGAGGTAAAGCAATAGCCAAATCATCTATTGCTG 736
 Qy 481 AGAGCACAGCGGAGGAGCAAGGATCGGATATAAAACCCAGGCAATTCGAGCCGCAACGG 540
 Db 737 AGAGCACAGCGGAGGAGCAAGGATCGGATATAAAACCCAGGCAATTCGAGCCGCAACAG 796
 Qy 541 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTCTACTCTATTGCTACTCT 600
 Db 797 CAACCCCTTTGGGTCCCTCCCTTTGATGGGGCTCTGTTTCTACTCTATTGCTACTCT 856
 Qy 601 ATTAATCTTGCACCTGAA 619
 Db 857 ATTAATCATGCAACTGCA 875

RESULT 8

US-10-087-192-910
 ; Sequence 910, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 910
 ; LENGTH: 21646
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(21646)
 ; OTHER INFORMATION: n = A, T, C or G
 ; US-10-087-192-910

Query Match 85.6%; Score 543.4; DB 13; Length 21646;
 Best Local Similarity 94.8%; Pred. No. 5e-156;
 Matches 562; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
 Qy 1 CCCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTTCCAGAACTCAAACTGTAAACTA 60
 Db 19571 CCCTGTATCTTTAACTCTCTTTGTTAAAGTTTGTCTTCCAGAACTGTAAACTA 19630
 Qy 61 CAATTTGTTCTTCAATGGAGCACAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 120
 Db 19631 CAATTTGTTCTTCAATGGAGCCCAAGATGAGTCCATGACTTAAGATCCACCGACCC 19690
 Qy 121 CTGAGCCGCTGTAGCCATGCTCCGATGTTAATGACATTAAGAGCAACCCCTCCCGAG 180
 Db 19691 CTGAGCCGCTGTAGCCATGCTCCGATGTTAATGACATTAAGAGCAACCCCTCCCGAG 19750
 Qy 181 GAAATCTCAATGCAACCCCTCTATATGCCCAATTCAGCGGAGCAAGTTAGAGCGGT 240
 Db 19751 GAAATCTCAATGCAACCCCTCTATATGCCCAATTCAGCGGAGCAAGTTAGAGCGGT 19810
 Qy 241 CATCAGCAACCTCCCAACAGCACTTGGTGTTCCTGTTGAGAGGGGAGCTGAGAGAC 300
 Db 19811 CATCAGCAACCTCCCAACAGCACTTGGTGTTCCTGTTGAGAGGGGAGCTGAGAGAC 19870
 Qy 301 AGGACTAGCTGATTTCTTAGGCCAAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 19871 AAGACTAGCTGATTTCTTAGGCCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 19930
 Qy 361 GCATCCACCTTAACATAGGGCTTGCACCTTAGCTCACACCCGACCAATCAGAGAGCTC 420

QY 361 GCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCGACCAATCAGAGAGCTC 420
 Db 2671 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 2730
 QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
 Db 2731 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 2790
 QY 481 AGAGCAGCGGAGGAGGACAGATCGGATATATAACCCAGGCAATTCGAGCCGGCAACGG 540
 Db 2791 AGAGCAGCAGGAGGAGCAATGATCGGATATATAACCCAAAGTCTTCGAGCCGGCAACGG 2850
 QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGCGCTCTCTTTTCACTCTATTTCACCTCT 600
 Db 2851 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 2910
 QY 601 ATTAAATCTTGCAACTGAAAAAATAAAAAA 635
 Db 2911 ATTAAATCTTGCAACTGAAAAAATAAAAAA 2945

RESULT 10
 US-10-016-249-3
 ; Sequence 3, Application US/10016249
 ; Publication No. US20030100053A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: LaVallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Mi, Sha
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6006B A1172A
 ; CURRENT APPLICATION NUMBER: US/10/016,249
 ; CURRENT FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: US/09/175,928
 ; PRIOR FILING DATE: 1998-10-20
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2946
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-016-249-3

Query Match 84.1%; Score 533.8; DB 15; Length 2946;
 Best Local Similarity 91.5%; Pred. No. 1.9e-155;
 Matches 581; Conservative 0; Mismatches 42; Indels 12; Gaps 1;
 QY 1 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAACTA 60
 Db 2323 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAATCGAAGTGTAAACTA 2382
 QY 61 CAAATTGTTCTTCAAAATGGAGCACCAAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 2383 -----CAAAATGGAGCCCAAGATGCAAGTCCAAAGACTAAGATCTACCGGACACCC 2430
 QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 180
 Db 2431 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 2490
 QY 181 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 Db 2491 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 2550
 QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
 Db 2551 CGTCGGCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGATGGGGACTGAGAGAC 2610

QY 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGAAAGGTGACT 360
 Db 2611 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGAAAGGTGACC 2670
 QY 361 GATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCGGACCAATCAGAGAGCTC 420
 Db 2671 ACATCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 2730
 QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 480
 Db 2731 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCCAAATCATCTATTGCCTG 2790
 QY 481 AGAGCAGCGGAGGAGCAAGGATCGGATATATAACCCAGGCAATTCGAGCCGGCAACGG 540
 Db 2791 AGAGCAGCAGGAGGAGCAATGATCGGATATATAACCCAAAGTCTTCGAGCCGGCAACGG 2850
 QY 541 CAACCCCTTTGGTCCCTCCCTTTGTATGGGCGCTCTCTTTTCACTCTATTTCACCTCT 600
 Db 2851 CAACCCCTTTGGTCCCTCCCTTTGTATGGGAGCTCTCTTTTCACTCTATTTCACCTCT 2910
 QY 601 ATTAAATCTTGCAACTGAAAAAATAAAAAA 635
 Db 2911 ATTAAATCTTGCAACTGAAAAAATAAAAAA 2945

RESULT 11
 US-09-731-231A-3/c
 ; Sequence 3, Application US/09731231A
 ; Patent No. US20020082189A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GUEGLER, Karl et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: C1001007
 ; CURRENT APPLICATION NUMBER: US/09/731,231A
 ; CURRENT FILING DATE: 2000-12-07
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 326014
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(326014)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-731-231A-3

Query Match 83.9%; Score 532.8; DB 9; Length 326014;
 Best Local Similarity 90.5%; Pred. No. 3.6e-154;
 Matches 583; Conservative 0; Mismatches 52; Indels 9; Gaps 1;
 QY 1 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAACTA 60
 Db 170822 CCCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAAATTTGAAGCTCTAAACTA 170763
 QY 61 CAAATTGTTCTTCAAAATGGAGCACCAAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 Db 170762 CAAATTGTTCTTCAAAATGGAGCCCAAGATGCAAGTCTACCGGACACCC 170703
 QY 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG 180
 Db 170702 CTGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCAGAG 170643
 QY 181 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
 Db 170642 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 170583
 QY 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
 Db 170582 CATCGGCCAATCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 170523

QY	301	AGGACTAGCTGGATTTCCTAGGCCAAACGAAGAAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACT	360
Db	170522	AGGACTAGCTGGATTTCCTAGGCCAAACGAAGAAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACT	170463
QY	361	GCATCCACCTCTAAACATGGGGCTTGCACCTTAGCTACACCCGACCAATC-----	411
Db	170462	GCTTCTACCTTTAAACCGGGCTTGCACCTTAGCTACACCTTAGCTAGCAATCAGGTAGGAA	170403
QY	412	AGAGAGCTCACTAAATGCTAAATAGGCACAAATAGGAGTAAAGAAATAGCCAAATCATC	471
Db	170402	AGAGAGCTCACTAAATGCTAAATAGGCACAAATAGGAGTAAAGAAATAGCCAAATCATC	170343
QY	472	TATTCCTCTGAGACACAGCGGAGGACAAAGGATCGGGATATATAACCCAGGCAATTCGAGC	531
Db	170342	TATTCCTCTGAGACACAGCGGAGGACAAAGGATCGGGATATATAACCCAGGCAATTCGAGC	170283
QY	532	CGGCAACGGCAACCCCTTTGGGTCCCTCTTTGATGGGCGCTCTGTTTCACTCTA	591
Db	170282	CAGCAATGGCTACCATTTTGGGTCCCTCTTTGATGGGCGCTCTGTTTCACTCTA	170223
QY	592	TTTCACTCTATTAAATCTTGCACCTGAAAAAATAAAAAA	635
Db	170222	TTTCACTCTATTAAATCTTGCACAGCAAAAAAATAAAAAA	170179

RESULT 13
US-10-632-793-25
; Sequence 25, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-25

Query Match	82.8%;	Score 525.8;	DB 17;	Length 1136;
Best Local Similarity	90.7%;	Pred. No. 3.9e-153;		
Matches	576;	Conservative	0;	Mismatches 45;
			Indels	12;
			Gaps	1;
QY	1	CCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCAGAAATCAAAACTGTAAACTA	60	
Db	510	CCCTGTATCTTTAACTCTCTTGTAACTTTGTCTCTTCAGAAATCGAAGCTGTAAACTA	569	
QY	61	CAAAATTTGTTCTTCAATGGAGCACAGATGGAGTCCATGATCAAGATCCACCGTGACCC	120	
Db	570	-----CAAAATGGAGCCCAAGATGCAGTCAAGACTTAAGATCTACCGCAGACCC	617	
QY	121	CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTAAGAGGACCCCTCCCGAG	180	
Db	618	CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTAAGAGGACCCCTCCCGAG	677	
QY	181	GAATCTCAACTGCACAAACCCCTCTATATGCCCAATTCAGCGGAGACGTTAGACGGT	240	
Db	678	GAATCTCAAGTGCACAACTCTACTACGCCCAATTCAGCGGAGACGTTAGACGGT	737	

QY	301	AGGACTAGCTGGATTTCCTAGGCCAAACGAAGAAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACT	360
Db	170522	AGGACTAGCTGGATTTCCTAGGCCAAACGAAGAAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACT	170463
QY	361	GCATCCACCTCTAAACATGGGGCTTGCACCTTAGCTACACCCGACCAATC-----	411
Db	170462	GCTTCTACCTTTAAACCGGGCTTGCACCTTAGCTACACCTTAGCTAGCAATCAGGTAGGAA	170403
QY	412	AGAGAGCTCACTAAATGCTAAATAGGCACAAATAGGAGTAAAGAAATAGCCAAATCATC	471
Db	170402	AGAGAGCTCACTAAATGCTAAATAGGCACAAATAGGAGTAAAGAAATAGCCAAATCATC	170343
QY	472	TATTCCTCTGAGACACAGCGGAGGACAAAGGATCGGGATATATAACCCAGGCAATTCGAGC	531
Db	170342	TATTCCTCTGAGACACAGCGGAGGACAAAGGATCGGGATATATAACCCAGGCAATTCGAGC	170283
QY	532	CGGCAACGGCAACCCCTTTGGGTCCCTCTTTGATGGGCGCTCTGTTTCACTCTA	591
Db	170282	CAGCAATGGCTACCATTTTGGGTCCCTCTTTGATGGGCGCTCTGTTTCACTCTA	170223
QY	592	TTTCACTCTATTAAATCTTGCACCTGAAAAAATAAAAAA	635
Db	170222	TTTCACTCTATTAAATCTTGCACAGCAAAAAAATAAAAAA	170179

RESULT 12
US-10-751-985-3/C
; Sequence 3, Application US/10751985
; Publication No. US20040126861A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001007CON
; CURRENT APPLICATION NUMBER: US/10/751,985
; CURRENT FILING DATE: 2004-01-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 326014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(326014)
; OTHER INFORMATION: n = A, T, C or G
US-10-751-985-3

Query Match	83.9%;	Score 532.8;	DB 18;	Length 326014;
Best Local Similarity	90.5%;	Pred. No. 3.6e-154;		
Matches	583;	Conservative	0;	Mismatches 52;
			Indels	9;
			Gaps	1;
QY	1	CCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCAGAAATCAAAACTGTAAACTA	60	
Db	170822	CCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCAGAAATGAAGCTCTAAACTA	170763	
QY	61	CAAAATTTGTTCTTCAATGGAGCACAGATGGAGTCCATGATCAAGATCCACCGTGACCC	120	
Db	170762	CAAAATTTGTTCTTCAATGGAGCCCAAGCAGCTCCATGATCAAGATCTACCGCAGACCC	170703	
QY	121	CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTAAGAGCACCCTCCCGAG	180	
Db	170702	CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTCGAAGGACCCCTCCAGAG	170643	
QY	181	GAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGACGTTAGACGGT	240	
Db	170642	GAATCTCAACTGCATRAACCCCACTACGCCCAATTCAGCGGAGACGTTAGACGAGT	170583	
QY	241	CATCAGCCAACTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC	300	
Db	170582	CATCGGCCATCTCCCAACAGCACTTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC	170523	

241	Qy	CATCAGCCAACTCTCCCAAAGACACTTTGGGTTTTCTGTTGAGAGGGGGGACCTGAGAC	300
738	Db	CGTCGGCCAACTCTCCCAAAGACACTTTAGGTCTTTCTGTGTTGAGATGGGGACTGAGAGAC	797
301	Qy	AGGACTAGCTGGATTTCTTAGGCCAAAGAGAATCCCTAAGCCCTAGCTGGGAAGGTGACT	360
798	Db	AGGACTAGCTGGATTTCTTAGCTGACTAAGATCCCTAAGCCCTAGCTGGGAAGGTGACC	857
361	Qy	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACCCGACCAATCAGAGAGCTC	420
858	Db	ACATCCACCTTTAAACACAGGGGCTTGCAACTTAGTTCACACCTGACCAATCAGAGAGCTC	917
421	Qy	ACTAAATGCTAATTAGGCGAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCGCTG	480
918	Db	ACTAAATGCTAATTAGGCGAAAGACAGGAGGTAAAGAAATAGCCAATCATCTATTGCGATG	977
481	Qy	AGAGCACAGCGGGAGGAGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGGCAACGG	540
978	Db	AGAGCACAGCAGGAGGAGCAATGATCGGGATATAAACCAAGCTTCGAGCCGGCAACGG	1037
541	Qy	CAACCCCTTTGGGTCCCTCCCTTTGTATGGCGCTCTGTTTTCACTCTATTTCACCTCT	600
1038	Db	CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTTCACTGATTTCACCTCT	1097
601	Qy	ATTAAATCTTGCAACTGAAAAAATAAAAAA	635
1098	Db	ATTAAATCTTGCACTCGAAAAAATAAAAAA	1132

RESULT 14
US-10-719-993-7065
; Sequence 7065, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7065
; LENGTH: 283351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(283351)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tab
US-10-719-993-7065

Db	175045	TGGGTTTTCTGTTGAGAGGGGACCTGAGAGCAGGACTAGCTGGATTTCTTAGGCCGA	175104
Qy	327	CGAAGAACTCCTAAGCCTAGCTGGGAAGGTGACTGCACTCCACCTTAAACATGGGGCTTG	386
Db	175105	CTAAGAACTCCTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG	175164
Qy	387	CNACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAATGCTAAATTAGGCCAAAAATA	446
Db	175165	CACTTTAGCTCACAC-----CCAACGAGAGGTTCACTAAAATGCTAAATTAGGCCAAAAATA	175220
Qy	447	GGAGTTAAAGAAATAGCCAATCATCTATTGCTCGAGACACAGCGGGAGGGAACAAGGATC	506
Db	175221	GGAGGTACAGAAATAGCCAATCATCTATTGCTCGAGACACAGCGGGAGGGAACAAGGATC	175280
Qy	507	GGGATATAACCCAGGCAATTCGAGCGGGCAACGCCCAACCCCTTTGGGTCCCTCCCTTT	566
Db	175281	AGGATATAAACCCAGGCAATTCGAGCTGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT	175340
Qy	567	GTATGGGGCTCTGTTTTCACTCTATTATTTCACTCTATTAAATCTTGCAACTGCAAAAAAAA	626
Db	175341	GTATGGAGCTCTGTTTTCA--CTGTTTCACTCTATTAAATCTTGCAACTGCAAAAAAGA	175398
Qy	627	AAAAAAA 634	
Db	175399	AAAAAAA 175406	

```

RESULT 15
US-10-741-600-17995
; Sequence 17995, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17995
; LENGTH: 283351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(283351)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Table
US-10-741-600-17995

```

Db 175045 TGGGTTTCTGTTGAGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGA 175104
Qy 327 CGAAGAATCCCTTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTCTAAACATGGGGCTTG 386
Db 175105 CTAAGAATCCCTTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 175164
Qy 387 CAACTTAGCTCACACCGGCAATCAGAGAGCTCACTAAATGCTAAATTAGGCAAAATA 446
Db 175165 CAACTTAGCTCACAC-----CCAACAGAGAGTTTCACTAAATGCTAAATTAGGCAAAATA 175220
Qy 447 GGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGACACGCGGAGGACAGAGATC 506
Db 175221 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGAGAGACACGCGGAGGACAGAGATC 175280
Qy 507 GGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 566
Db 175281 AGGATATAAACCCAGGCAATTCGAGCTGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 175340
Qy 567 GTATGGGCGCTGTGTTTCACTCTATTTCACCTCTATTAATCTTCAACTGAAAAAATA 626
Db 175341 GTATGGGAGCTGTGTTTCA--CTGTTTCACTCTATTAATCTTGCACCTGCAAAATA 175398
Qy 627 AAAAAA 634
Db 175399 AAAAAA 175406

RESULT 16
US-10-719-993-6815
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815

Query Match 82.3%; Score 522.8; DB 18; Length 1980090;
Best Local Similarity 93.4%; Pred. No. 1.1e-150;
Matches 568; Conservative 1; Mismatches 33; Indels 6; Gaps 2;
Qy 27 GTTTGTCTCTCCAGAAATCAAACTGTAAACTACAAATGTGTTCTTCAAATGGAGCACCA 86
Db 1777083 GTTTGTCTCTCCAGAAATCAAACTGTAAACTACAAATGTGTTCTTCAAATGGAGCCCCA 1777142
Qy 87 GATGGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGCTGCTAGCCATGCTC 146
Db 1777143 GATGCAGTCCATGACTAAGATCCACCGACAGCCCTGGACGCGCTGCTAGCCATGCTC 1777202
Qy 147 CGATGTTAATGACATTTGAAGGCAACCCCTCCGAGGAAATCTCAACTGCAACACCTACT 206
Db 1777203 CAATGTTAATGACATTTGAAGGCAACCCCTCCGAGGAAATCTCAACTGCAACACCTACT 1777262
Qy 207 ATGCCCAATTCAGCGGGAAGCTTTAGAGCGGTCTATCAGCCAACTCTCCCAACAGCACT 266
Db 1777263 ATGCCCAATTCAGCGGGAAGCTTTAGAGCAGTCTATGGCCAACTCTCCCAACAGCACT 1777322
Qy 267 TGGGTTTTCTGTTGAGAGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAA 326
Db 1777323 TGGGTTTTCTGTTGAGAGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGA 1777382

Qy 327 CGAAGAATCCCTTAAGCCTAGCTGGGAAGGTGACTGCATCCACCTCTAAACATGGGGCTTG 386
Db 1777383 CTAAGAATCCCTTAAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 1777442
Qy 387 CAACTTAGCTCACACCGGCAATCAGAGAGCTCACTAAATGCTAAATTAGGCAAAATA 446
Db 1777443 CAACTTAGCTCACAC-----CCAACAGAGAGTTTCACTAAATGCTAAATTAGGCAAAATA 1777498
Qy 447 GGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGACACGCGGAGGACAGAGATC 506
Db 1777499 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGAGAGACACGCGGAGGACAGAGATC 1777558
Qy 507 GGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 566
Db 1777559 AGGATATAAACCCAGGCAATTCGAGCTGGCAACGGCAACCCCTTTGGGTCCCTCCCTTT 1777618
Qy 567 GTATGGGCGCTGTGTTTCACTCTATTTCACCTCTATTAATCTTGCACCTGAAAAAATA 626
Db 1777619 GTATGGGAGCTGTGTTTCA--CTGTTTCACTCTATTAATCTTGCACCTGCAAAATA 1777676
Qy 627 AAAAAA 634
Db 1777677 AAAAAA 1777684

RESULT 17
US-10-741-600-17676
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17676
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

Query Match 82.3%; Score 522.8; DB 19; Length 1980090;
Best Local Similarity 93.4%; Pred. No. 1.1e-150;
Matches 568; Conservative 1; Mismatches 33; Indels 6; Gaps 2;
Qy 27 GTTTGTCTCTCCAGAAATCAAACTGTAAACTACAAATGTGTTCTTCAAATGGAGCACCA 86
Db 1777083 GTTTGTCTCTCCAGAAATGTAAACTGTAAACTGTAAACTGTGTTCTTCAAATGGAGCCCCA 1777142
Qy 87 GATGGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGCTGCTAGCCATGCTC 146
Db 1777143 GATGCAGTCCATGACTAAGATCCACCGACAGCCCTGGACGCGCTGCTAGCCATGCTC 1777202
Qy 147 CGATGTTAATGACATTTGAAGGCAACCCCTCCGAGGAAATCTCAACTGCAACACCTACT 206
Db 1777203 CAATGTTAATGACATTTGAAGGCAACCCCTCCGAGGAAATCTCAACTGCAACACCTACT 1777262
Qy 207 ATGCCCAATTCAGCGGGAAGCTTTAGAGCGGTCTATCAGCCAACTCTCCCAACAGCACT 266
Db 1777263 ATGCCCAATTCAGCGGGAAGCTTTAGAGCAGTCTATGGCCAACTCTCCCAACAGCACT 1777322
Qy 267 TGGGTTTTCTGTTGAGAGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAA 326
Db 1777323 TGGGTTTTCTGTTGAGAGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGA 1777382

Qy 327 CGAAGATCCCTAGCCTAGCTGGGAAGGTGACTGATCCACCTCTAAACATGGGGCTTG 386
Db 1777383 CTAAGAATCCCTAGCCTAGCTGGGAAGGTGACCGCATCCACCTTTAAACACGGGGCTTG 1777442
Qy 387 CAATCTAGCTCACACCCGACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAATA 446
Db 1777443 CAATCTAGCTCACAC-----CCAACGAGAGTTCACCTAAAATGCTAATAGGCAAAATA 1777498
Qy 447 GGAGGTAAGAAATAGCAATCATCTATTGCTTGAGAGCACAGGGGAGGACAGAGATC 506
Db 1777499 GGAGGTACAGAAATAGCAATCATCTATTGCTTGAGAGCACAGGGGAGGACAGAGATC 1777558
Qy 507 GGGATATAAACCCAGGCAATCGAGCCGCAACCGCAACCCCTTTGGGTCCCTCCCTTT 566
Db 1777559 AGGATATAAACCAAGGCAATCGAGCTGGCAACCGCAACCCCTTTGGGTCCCTCCCTTT 1777618
Qy 567 GTATGGCGCTCTCTTTTCACTCTATTCTCTATTCTATTAAATCTTTGCAACTGAAAAAAA 626
Db 1777619 GTATGGGAGCTCTGTTTCA--CTGTTTCACTCTATTAAATCTTTGCAACTGCAAAAAAGA 1777676
Qy 627 AAAAAAA 634
Db 1777677 AAAAAAA 1777684

RESULT 18

US-10-632-793-26
; Sequence 26, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-26

Query Match 82.3%; Score 522.6; DB 17; Length 2782;
Best Local Similarity 90.4%; Pred. No. 5.9e-152;
Matches 574; Conservative 0; Mismatches 49; Indels 12; Gaps 1;
Qy 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTTGTCTCTCCAGAACTCAAACTGTAATACTA 60
Db 2158 CCCTGTATCTTTGACCTCTTTGTTAACTTTGTCTCTCCAGAACTGTAATACTA 2217
Qy 61 CAATTTGTTCTTCAATGGAGCACAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 120
Db 2218 -----CAATGGAGCCCAAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 2265
Qy 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGATTAAGGCAACCCCTCCCGAG 180
Db 2266 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGATTAAGGCAACCCCTCCCGAG 2325
Qy 181 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGAGGAGTGTAGAGCGGT 240
Db 2326 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGAGGAGTGTAGAGCGGT 2385
Qy 241 CATCAGCCAACTCCCCCAACAGCACTTGGGTTTTCTCTGTGAGAGGGGGGACTGAGAGAC 300

Db 2386 GGTCCGCCAACCTCCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGACTGAGAGAC 2445
Qy 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAAATCCTTAAGCCTAGCTGGGAAGGTGACT 360
Db 2446 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAAATCCTTAAGCCTAGCTGGGAAGGTGACC 2505
Qy 361 GCATCCACCTCTAAACATGGGGCTTCAACTTAGCTCACACCCGACCAATCAGAGAGCTC 420
Db 2506 ACATCCACCTTTAAACACAGGGGCTTCAACTTAGCTCACACCTGACCAATCAGAGAGCTC 2565
Qy 421 ACTAAATGCTAATTAGGCAAAATAGGAGTTAAAGAAATAGCAATCATCTATTGCGCTG 480
Db 2566 ACTAAATGCTAATTAGGCAAAACAGAGGAGTTAAAGAAATAGCAATCATCTATTGCGCTG 2625
Qy 481 AGAGCAGAGCGGAGGAGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGCAACGG 540
Db 2626 AGAGCAGAGCGAGGAGCAATGATCGGGATATAAACCCAGGTTTTCGAGCCGCAACGG 2685
Qy 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCACCTCT 600
Db 2686 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCTCT 2745
Qy 601 ATTAATCTTGCACCTGAAAAAATAAAAAAAAAAAAA 635
Db 2746 ATTAATCTTGCACCTGCAAAAAAATAAAAAAAAAAAAA 2780

RESULT 19

US-10-133-036-1
; Sequence 1, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
; APPLICANT: Conrad, Bernard
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Human endogenous retrovirus
US-10-133-036-1

Query Match 82.0%; Score 521; DB 17; Length 2782;
Best Local Similarity 90.2%; Pred. No. 1.9e-151;
Matches 573; Conservative 0; Mismatches 50; Indels 12; Gaps 1;
Qy 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTTGTCTCTCCAGAACTCAAACTGTAATACTA 60
Db 2158 CCCTGTATCTTTGACCTCTTTGTTAACTTTGTCTCTCCAGAACTGTAATACTA 2217
Qy 61 CAATTTGTTCTTCAATGGAGCACAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 120
Db 2218 -----CAATGGAGCCCAAGATGAGTCCATGACTTAAGATCCACCGTGGACCC 2265
Qy 121 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGATTAAGGCAACCCCTCCCGAG 180
Db 2266 CTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGATTAAGGCAACCCCTCCCGAG 2325
Qy 181 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGAGGAGTGTAGAGCGGT 240
Db 2326 GAAATCTCAATGCAACACCCCTACTATGCCCAATTCAGCGGAGGAGTGTAGAGCGGT 2385
Qy 241 CATCAGCCAACTCCCCCAACAGCACTTGGGTTTTCTCTGTGAGAGGGGGGACTGAGAGAC 300
Db 2386 GGTCCGCCAACCTCCCCCAACAGCACTTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 2445

301	AGGACTAGCTGGATTTCCTAGGCCAAACGAAGAATCCCTTAAGCCTAGCTGGGAGGTGACT	360
2446	AGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCTTTAAGCCTAGGTGGGAAGGTGACC	2505
361	GCATCCACCTCTAAACATGCGGCTTGCAACTTAGTCTCACCCGACCAATCAGAGAGCTC	420
2506	ACATCCACCTCTTTAAACGCGGCTTGCAACTTAGTCTCACCTGACCAATCAGAGAGCTC	2565
421	ACTAAATATGCTAATTAGGCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTG	480
2566	ACTAAATATGCTAATTAGGCAAAAGACAGGAGGTAAAGAAATAGCCAATCATTTATTGCCTG	2625
481	AGAGCACAGCGGAGGAGGACNAGNATCGGATATAAACCCAGGCAATCGAGCCGCGCAACGG	540
2626	AGAGCACAGCAGGAGGACAATGATCGGATATAAACCCAGTTTTCGAGCCGCGCAACGG	2685
541	CAACCCCTTTGGGTCCCTTCTGATGGGCGCTCTGTTTTCACCTCTAATTTCACTCT	600
2686	CAACCCCTTTGGGTCCCTTCTGATGGGAGCTCTGTTTTCACGCTATTTCACTCT	2745
601	ATTAAATCTTGCAACTGAAAAAATAAAAAA	635
2746	ATTAAATCTTGCAACTGCAAAAAAATAAAAAA	2780

```

RESULT 20
US-09-902-535-1
; Sequence 1, Application US/0902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: Mi, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preeclampsia and gestational trophoblast
; TITLE OF INVENTION: disorders
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930) ... (2546)
US-09-902-535-1

```

2553	Db	 COTCGGCCAACCTCCCAACAGCACCTTAGGCTTTCTGTTGAGATCGGGCACTGAGAGAC	2612
301	Qy	 AGGACTAGCTCGATTTCTTAGGGCCAAACGAAGAATCCCTAAGCCTAGCTGGGAAGTGTACT	360
2613	Db	 AGGACTAGCTCGATTTCTTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGTGTACC	2672
361	Qy	 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCGGACCAATCAGAGAGCTC	420
2673	Db	 ACATCCAGCTTTAAACACGGGGCTTGCACTTAGCTCACACCTGACCAATCAGAGAGCTC	2732
421	Qy	 ACTAAATGCTAAATTTAGGCAAAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCGCTG	480
2733	Db	 ACTAAATGCTAAATTTAGGCAAAAAACAGAGGTAAAGAAATAGCAATCATCTATTGCGCTG	2792
481	Qy	 AGAGCACACGGGGAGGGCAAGGATCGGGATATAAACCCAGGCAATTCGAGCCGGCAACGG	540
2793	Db	 AGAGCACACGACGAGGGGCAATGATCGGGATATAAACCCAGGCTTCGAGCCGGCAACGG	2852
541	Qy	 CAACCCCTTTGGGTCGCCCTCCCTTTGATGGGCGCTCTGTTTTCACCTCTATTTCACCTCT	600
2853	Db	 CAACCCCTTTGGGTCGCCCTCCCTTTGATGGGAGCTCTGTTTTTCATGCTATTTCACCTCT	2912
601	Qy	 ATTAAATCTTGCAACTG	617
2913	Db	 ATTAAATCTTGCAACTG	2929

```

RESULT 21
US-09-873-367C-81
; Sequence 81, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-81

```

	Query Match	81.3%	Score 516.2	DB 10	Length 56093
Best Local Similarity	91.1%				
Matches 564	Conservative 0	Mismatches 43	Indels 12	Gaps 1	
Qy	1	CCCTGTATCTTTAACTTCCTTTGTTAAAGTTTGCTCTTC	CAGAAATCAAAATCTTAAACTA	60	
Db	37274	CCCTGTATCTTTAACTTCCTTTGTTAACTTTGCTCTTC	CAGAAATCGAAGCTTGAANAATA	37333	
Qy	61	CAAAATGTTCTTTCAAATGGAGACACAGATGGAGTCCAT	GACTAAAGATCCACCGTGGAGCC	120	
Db	37334	-----CAAAATGGAGCCCAAGATGCAGCTTCCAAGACT	TAAGATCTTACCGCGAGACC	37381	

```

Db 37382 CTGGACGGCTCTAGCCACGATCTGATGTTAATGACATCAAGACGACCCCTCTGAG 37441
Qy 181 GAAATCTCAACTGACACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 37442 GAAATCTCAGCTGACACACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 37501
Qy 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
Db 37502 CTGGGCAACCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGATGGGAGCTGAGAGAC 37561
Qy 301 AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACT 360
Db 37562 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGCCTAGCTAGCTGGGAAGGTGACC 37621
Qy 361 GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCAGACCCGACCAATCAGAGAGCTC 420
Db 37622 ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCAGACCCGACCAATCAGAGAGCTC 37681
Qy 421 ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG 480
Db 37682 ACTAAATGCTAATTAGGCAAAAGAGAGGATTAAGAAATAGCAATCATCTATTGCTG 37741
Qy 481 AGAGCAGCGGGAGGAGCAAGGATCGGGATATAAAACCCAGGATTCGAGCCGGCAACGG 540
Db 37742 AGAGCAGCAGGAGGAGCAATGATCGGATATAAAACCCAGGCTTCGAGCCGGCAACGG 37801
Qy 541 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCT 600
Db 37802 CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCT 37861
Qy 601 ATTAATCTTGCACCTGAA 619
Db 37862 ATTAATCTTGCACCTGCA 37880

```

RESULT 22

```

US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauca
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT FILING DATE: 2003-08-04
; PRIOR FILING DATE: 2001-10-22
; PRIOR FILING DATE: 2000-01-21
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
; NAME/KEY: misc feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u

```

Query Match 80.5%; Score 511.4; DB 17; Length 7582;
Best Local Similarity 89.5%; Pred. No. 3e-148;
Matches 552; Conservative 11; Mismatches 42; Indels 12; Gaps 1;

Qy	1	CCCTGTATCTTTAACTCTCTTGTAACTTCTCTCCAGAAATCAAACTGTAAACTA	60
Db	6976	CCCTGTATCTTTTAACTCTCTTGTAACTTCTCTCCAGAAATCAAACTGTAAACTA	7035
Qy	61	CAAAATGTTCTTCAAAATGGAGACACAGATGGAGTCCATGATCAAGATCCACCGTGGACCC	120
Db	7036	-----CAAAATGGAGCCAGATGCACTCCAAGACTAAGATCTACCGACAGACCC	7083
Qy	121	CTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG	180
Db	7084	CTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAG	7143
Qy	181	GAAATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGAGAGCAGTTAGAGCGGT	240
Db	7144	GAAATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGAGAGCAGTTAGAGCGGT	7203
Qy	241	CATCAGCAACCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGAGCTGAGAGAC	300
Db	7204	SGTCGGCAACCTCCCAACAGCAGCTTGGGTTTTCTGTTGAGATGGGAGCTGAGAGAC	7263
Qy	301	AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT	360
Db	7264	AGGACTAGCTGGATTTCTTAGGCTGATTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACC	7323
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCAGACCCGACCAATCAGAGAGCTC	420
Db	7324	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCAGACCCGACCAATCAGAGAGCTC	7383
Qy	421	ACTAAATGCTAATTAGGCAAAATAGGAGTAAAGAAATAGCAATCATCTATTGCTG	480
Db	7384	ACTAAATGCTAATTAGGCAAAAGAGAGTAAAGAAATAGCAATCATCTATTGCTG	7443
Qy	481	AGAGCAGCGGAGGAGCAAGGATCGGGATATAAAACCCAGGATTCGAGCCGGCAACGG	540
Db	7444	AGAGCAGCAGGAGGAGCAATGATCGGGATATAAAACCCAGGATTCGAGCCGGCAACGG	7503
Qy	541	CAACCCCTTTGGGTCCCTCCCTTTGTATGGGGCTCTGTTTCACTCTATTTCACCT	600
Db	7504	CAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCACTCTATTTCACCT	7563

Fri Feb 25 16:26:30 2005

```
Qy 601 ATTAATCTTGAACGTG 617
Db 7564 ATTAATCTTGCARCTG 7580

RESULT 23
US-10-087-192-730/c
; Sequence 730, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 730
; LENGTH: 161334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(161334)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-730

Query Match 80.4%; Score 510.6; DB 13; Length 161334;
Best Local Similarity 90.0%; Pred. No. 2,2e-147;
Matches 575; Conservative 0; Mismatches 49; Indels 15; Gaps 2;

Qy 1 CCTGTATCTTTAACTCTCTTGTAAAGTTTGCTCTCCAGAAATCAAACCTGTAAACTA 60
Db 100611 CCTGTATCTTTAACTCTCTTGTAAAGTTTGCTCTCCAGAAATCAAAGTTGTAAAGCTA 100552

Qy 61 CAAATGTTCTTCAATGAGCAGCAGATGGATCCATGACTTAAGATCCACCGTGACCC 120
Db 100551 CAAATGTTCTTCAATGAGCAGCAGATGGATCCATGACTTAAGATCCACCGTGACCC 100492

Qy 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCCACCCCTCCCGAG 180
Db 100491 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCCACCCCTCCCGAG 100432

Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCGCGGAGACGATTAGACGGT 240
Db 100431 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCGCGGAGACGATTAGACGGT 100372

Qy 241 CATCAGCCACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGGACTTGAGAGAC 300
Db 100371 CGTCAGCCACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGGACTTGAGAGAC 100312

Qy 301 AGGACTAGCTGGATTCTTAGGCCAACGAAAGAAATCCCTAAGCCTAGCTGGGAAAGTGACT 360
Db 100311 AGGATTAGCTGGATTCTTAGGCCAACGAAAGAAATCCCAAGCCTAGCTGGGAAAGTGACC 100252

Qy 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATC 411
Db 100251 ACATCCACCTCTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGGTAGTAA 100192

Qy 412 AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATC 471
Db 100191 AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATC 100132

Qy 472 TATTGCTGAGAGCAGCGGGAGGGAACAAGGATCGGGATATAAACCAGGCAATTCGAGC 531
Db 100131 TATCGCTGAGAGCAGCGGGAGGGAACAATGATCGGGATATAAACCAGGCAATTCGAGC 100072

Qy 532 CGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGCGCTCTGTTTTCACCTCTA 591
Db 100071 CGCAACGGCTACCTTCTTTGGGTCCCTCCCTTTGATGGAGCTCTCTCT-----GT 100018

Qy 592 TTTCACTCTATTAAATCTTGCAACTGAAAAAATA 630
Db 100017 CTTCACTCTATTAAATATTGCAACTGCAAAAAAATA 99979

RESULT 24
US-10-087-192-1666
; Sequence 1666, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1666
; LENGTH: 285020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(285020)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1666

Query Match 78.9%; Score 500.8; DB 13; Length 285020;
Best Local Similarity 89.3%; Pred. No. 3.3e-144;
Matches 577; Conservative 0; Mismatches 57; Indels 12; Gaps 3;

Qy 1 CCTGTATCTTTAACTCTCTTGTAAAGTTTGCTCTCCAGAAATCAAACCTGTAAACTA 60
Db 279486 CCTGTATCTTTAACTCTCTTGTAAAGTTTGCTCTCCAGAAATCGAAGCTGTAAACTA 279545

Qy 61 CAAATGTTCTTCAATGAGCAGCAGATGGATCCATGACTTAAGATCCACCGTGACCC 120
Db 279546 CAAATGTTCTTCAATGAGCAGCAGATGGATCCATGACTTAAGATCCACCGTGACCC 279605

Qy 121 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCCACCCCTCCCGAG 180
Db 279606 CTGGACCGGCTCTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCCACCCCTCCCGAG 279665

Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGACGATTAGACGGT 240
Db 279666 GAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGAGACGATTAGACGGT 279725

Qy 241 CATCAGCCACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGGACTTGAGAGAC 300
Db 279726 CATTCGCCAACCTCCCAACAGCAGCTTGGGTTTTCTGTGTAGAGGGGGGACTTGAGAGAC 279785

Qy 301 AGGACTAGCTGGATTCTTAGGCCAACGAAAG--AATCCCTAAGCCTAGCTGGGAAAGGTGA 358
Db 279786 AGGACTAGCTGGATTCTTAGGCCAACGAAAG--AATCCCTAAGCCTAGCTGGGAAAGGTGA 279845

Qy 359 CTGCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATC----- 411
Db 279846 CGGCATCCACCTCTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGGGAGT 279905

Qy 412 --AGAGAGCTCACTAAATGCTAAATAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCA 469
Db 279906 AAGAGAGCTCACTAAATGCTAAATAGGCTAAATAGGCTAAACAGAGGTTAAAGAAATAGCCAAATCA 279965
```


; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 678
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-678

Query Match
Best Local Similarity 77.7%; Score 493.2; DB 17; Length 2052;
Matches 549; Conservative 0; Mismatches 43; Indels 10; Gaps 2;

1 CCCTGTATCTTTAAACCTCTCTTGAAGTTTGTCTCTTCCAGAAATCAAACTGTAAACTA 60
Dbb 993 CCCTGTATCTTTAAACCTCTCTTGAAGTTTGTCTCTTCCAGAAATCAAAAGCTGTAAACAA 1052

61 CAAATTTCTTCAAATGGAGACCAAGATGGATGCTCATGATCAAGATCCACCGTGGACCC 120
Dbb 1053 CACATCGTTCTTCAAATGGAGACCAAGATGGATGCTCATGATCAAGATCTACCGGGATCC 1112

121 CTGACCGGCTCTAGCCCATGCTCGATGTTAATGACATTTGAAGCACCCTCCCGAG 180
Dbb 1113 CTGACCGGCTCTAGCCCATGCTCGATGTTAATGACATTTGAAGCACCCTCCCGAG 1172

181 GAAATCTCAATGCAACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Dbb 1173 GAAATCTCAATGCAACACCCCTACTATGCCCCCAATTTAAGCAGGAGCAGTTGAGCAGT 1232

241 CATCAGCGCACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGACTCAGAGAC 300
Dbb 1233 CGAGCGGCACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGACTCAGAGAC 1292

301 AGGACTAGCTGGATTTCTTAGGCCAAACAGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 360
Dbb 1293 AGGACTAGCTGGATTTCTTAGGCCAAACAGAAATCCCTAAGCCTTAGCTGGGAAGGTGACT 1352

361 GCATCCACCTCTAAACATGGGGCTTGCATTTAGCTTACACCCGACCAATC----- 411
Dbb 1353 GTATCCACCTTTAAACACAGGGCTTGCATTTAGCTTACACCCGACCAATCAGGTAGTAA 1412

412 AGAGAGCTCACTAAATGCTAATTTAGGCACAAATAGGAGGTAAAGAAATAGCCTAATCATC 471
Dbb 1413 AGAGAGCTCACTAAATGCTAATTTAGGCACAAATAGGAGGTAAAGAAATAGCCTAAT-ATC 1471

472 TATTCCCTGAGACACAGCGGAGGCAAGGATCGGATATAAACCCAGGCAATTCGAGC 531
Dbb 1472 TATCACCTGAGAGTACAGGGGAGGCAATGATTGGGATAGAAACCCAGGCAATTCGAGC 1531

532 CGGCAACGGCAACCCCTTTGGGTTCCCTTGTATGGCGCTCTGTTTTCACCTCTA 591
Dbb 1532 CGGCAACGGCAACCCCTTTGGGTTCCCTTGTATGGCGCTCTGTTTTCACCTCTA 1591

592 TT 593
Dbb 1592 TT 1593

RESULT 28
US-09-997-722-148
; Sequence 148, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02

; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-28

Query Match
Best Local Similarity 78.0%; Score 495.6; DB 17; Length 3372;
Matches 522; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

70 CTTCAAATGGAGCACAGATGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGG 129
Dbb 2802 CTACAAATGGAGCCCAAGATGAGTCCCAAGACTAAGATCTACCGCAGACCCCTGGACCGG 2861

130 CTTGCTAGCCCATGCTCCCATGTTAATGACATTTGAAGGCACCCCTCCCGAGGAAATCTCA 189
Dbb 2862 CTTGCTAGCCCATGCTCCCATGTTAATGACATTTGAAGGCACCCCTCCCGAGGAAATCTCA 2921

190 ACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTCAATCAGCA 249
Dbb 2922 GCTGCACAAACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTCAATCAGCA 2981

250 ACCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGACTGAGAGACAGGACTAGC 309
Dbb 2982 ACCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGACTGAGAGACAGGACTAGC 3041

310 TGGATTTCTTAGGCTGATTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACCAATCCACC 3101
Dbb 3042 TGGATTTCTTAGGCTGATTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGACCAATCCACC 3161

370 TCTAAACATGGGCTTGGCACTTAGCTCAGCCCGACCAATCAGAGAGCTCACTAAATG 429
Dbb 3102 TTTAAACACAGGGCTTGGCACTTAGCTCAGCCCGACCAATCAGAGAGCTCACTAAATG 3161

430 CTAATTTAGGCACAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGAGCAG 489
Dbb 3162 CTAATTTAGGCACAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTTCCTTGAGAGCAG 3221

490 CGGAGGGCAAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACCGGCAACCCCT 549
Dbb 3222 CAGGAGGGCAAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACCGGCAACCCCT 3281

550 TTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTTCACCTTATTTCACTTATTAATCT 609
Dbb 3282 TTGGGTCCTCCCTCTTGTATGGCGCTCTGTTTTCACCTTATTTCACTTATTAATCT 3341

610 TGCACCTGAAATGAAAAAAAAAAAAA 635
Dbb 3342 TGCACCTGCAAAAAAAAAAAAAA 3367

RESULT 27
US-10-276-774-678
; Sequence 678, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: NO. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914

; NUMBER OF SEQ ID NOS: 301
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 148
 ; LENGTH: 22436
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-997-722-148

Query Match 76.0%; Score 482.6; DB 11; Length 22436;

Best Local Similarity 87.6%; Pred. No. 4.8e-139; Indels 11; Gaps 3;
Matches 565; Conservative 0; Mismatches 69;

Qy 1 CCCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCCAGAACTCAAACTGTAAGAACTA 60

Db 5655 CCTGTACCTTTAACTCTCTTTAGTTTGTCTCTTCCAGAACTCAAACTGTAAGAACTA 5714

Qy 61 CAAATGTTCTTCAATGGAGCCAGCATGGAGTCCATGACTAAGATCCACCGTGACCC 120

Db 5715 CAAATGTTCTTCAATGGAGCCAGCATGGAGTCCATGACTAAGATCCACCGTGACCC 5774

Qy 121 CTGACCGGCTCTAGCC-CATGCTCCGATGTTAATGACATTAAGGCAACCCCTCCCGA 179

Db 5775 CTGACCTGGCTCTAGCCCTCTTCAATGTTAATGACATTAAGGCAACCTCTCCCGA 5834

Qy 180 GGAATCTCAACTGCACAACTCTATGCTCCCAATTCAGCGGAGCAGTGTAGAGCGG 239

Db 5835 GGAATCTCAACTGCACAACTCTATGCTCCCAATTCAGCGGAGCAGTGTAGAGCAA 5894

Qy 240 TCATGACCAACTCTCCCAACAGACATTTGGTCTTCTGTGAGAGGGGACTGAGAGA 299

Db 5895 TCATTGGCCAACTCTCCCAACAGACATTTGGTCTTCTGTGAGAGGGGACTGAGAG- 5953

Qy 300 CAGGACTAGCTGGATTTCTTAGCCACAGAGATCCCTAAGCCTAGCTGGAGGTGAC 359

Db 5954 CAGGACTAGCTGGATTTCTTAGGCTGACTAAGAACTCCCTAAGCCTAGCTGGAGGTGAC 6013

Qy 360 TGCATCCACTCTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 411

Db 6014 TGCATCCACTCTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAAGTAGTA 6073

Qy 412 -AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGAGGTTAAGAAATAGCAATCAT 470

Db 6074 AAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGAGGTTAAGAAATAGCAATCAT 6133

Qy 471 CTATTGCTGAGACACAGCGGAGGACAGAGTCCGGATATAACCCAGGCAATTCGAG 530

Db 6134 CTATCAGCTGAGACACAGCGGAGGACAGAGTCCGGATATAACCCAGGCAATTCGAG 6193

Qy 531 CCGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGTTGTTGGGGCTCTGTTTCACTCT 590

Db 6194 CCAGCAATGGCTACCTCTTTGGGTCCCTCCCTTTGTTGTTGGGGAGCTCTGTTTCACTCT 6253

Qy 591 ATTTCACCTATTAAATCTTGCAACTGAAATGAAATGAAATGAAATGAAATGAAAT 635

Db 6254 ATTAACCTTGCAACTGCAAAATATATATATATATATATATATATATATATATATAT 6298

RESULT 29

US-10-220-120-15

; Sequence 15, Application US/10220120

; Publication No. US20040048253A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: PANZER, Scott R.

; APPLICANT: SPIRO, Peter A.

; APPLICANT: BANVILLE, Steven C.

; APPLICANT: SHAH, Purvi

; APPLICANT: CHALUP, Michael S.

; APPLICANT: CHANG, Simon C.

; APPLICANT: CHEN, Alice

; APPLICANT: D'SA, Steven A.

; APPLICANT: AMSHEY, Stefan

; APPLICANT: DAHL, Christopher R.

; APPLICANT: DAM, Tam C.
 ; APPLICANT: DANIELS, Susan E.
 ; APPLICANT: DUFOUR, Gerard E.
 ; APPLICANT: FLORES, Vincent
 ; APPLICANT: FONG, Willy T.
 ; APPLICANT: GREENAWALT, Lila B.
 ; APPLICANT: HILLMAN, Jennifer L.
 ; APPLICANT: JONES, Anissa L.
 ; APPLICANT: LIU, Tommy F.
 ; APPLICANT: ROSEBERRY, Ann M.
 ; APPLICANT: ROSEN, Bruce H.
 ; APPLICANT: RUSSO, Frank D.
 ; APPLICANT: STOCKREHER, Theresa K.
 ; APPLICANT: DAFPO, Abel
 ; APPLICANT: WRIGHT, Rachel J.
 ; APPLICANT: YAP, Pierre E.
 ; APPLICANT: YU, Jimmy Y.
 ; APPLICANT: BRADLEY, Diana L.
 ; APPLICANT: BRATCHER, Shawn R.
 ; APPLICANT: CHEN, Wensheng
 ; APPLICANT: COHEN, Howard J.
 ; APPLICANT: HODGSON, David M.
 ; APPLICANT: LINCOLN, Stephen E.
 ; APPLICANT: JACKSON, Stuart
 ; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: PT-1113 PCT
 ; CURRENT APPLICATION NUMBER: US/10/220,120
 ; CURRENT FILING DATE: 2002-08-26
 ; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,776;
 ; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
 ; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
 ; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
 ; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
 ; 60/204,525; 60/205,285; 60/205,323; 60/205,287;
 ; 60/205,324; 60/205,286
 ; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
 ; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
 ; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
 ; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
 ; 2000-05-17; 2000-05-17; 2000-05-16; 2000-05-16; 2000-05-15;
 ; 2000-05-16; 2000-05-17; 2000-05-17; 2000-05-17; 2000-05-17;
 ; 2000-05-17; 2000-05-17
 ; NUMBER OF SEQ ID NOS: 422
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 15
 ; LENGTH: 849
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:429446.1:2000FEB18
 US-10-220-120-15

Query Match 75.8%; Score 481.4; DB 17; Length 849;

Best Local Similarity 89.1%; Pred. No. 2.5e-139;
Matches 534; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

Qy 27 GTTTGTCTCTCCAGAACTCAAACTGTAAACTTACAAATTTCTTCAAAATGGAGCACCA 86

Db 121 GTTTGTCTCTCCAGAACTCAAACTGTAAACTTACAAATTTCTTCAAAATGGAGCCCCA 180

Qy 87 GATGGAGTCCATGACTAAGATCCACCGTGGACCCCTTGACCGGCTCTAGCCCATGCTC 146

Db 181 GATGGAGTCCATGACTAAGATCCACCGTGGACCCCTTGACCGGCTCTAGCCCATGCTC 240

Qy 147 CGATGTTAATGACATTAAGGCAACCCCTCCGAGGAAATCTCACTGCAACACCCCTACT 206

Db 241 CGATGTTAATGACATTAAGGCAACCCCTCCGAGGAAATCTCACTGCAACACCCCTACT 300

Qy 207 ATGCCCCAATTCAGCGGAGGAGCAGTTAGAGCGGTTCATCAGCAACCTCCCCAACAGCACT 266

Db 301 ACGCCCAATTCAGCGGAGGAGCAGTTAGAGCGGTTCATCAGCAACCTCCCCAACAGCACT 360

Fri Feb 25 16:26:30 2005

QY 267 TGGGTTTTCTGTTGAGAGGGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAA 326
DB 361 TGAGTTTTCTGTTGAGAGGGGAGAACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAA 420
QY 327 CGAAGATCCCTAGCCTAGCTGGGAAGGTGACCTGATCCACCTCTAAACATGGGGTTG 386
DB 421 CTAAGATCCCTAAGCCTATCTGGGAAGGTGACCGTATCCACCTCTTAAACATAGGGTTG 480
QY 387 CAACTTAGCTCACCCGACCAATCA-----GAGAGCTCACTAAATGCTAAATTAG 437
DB 481 CAACTTAGCTCACCCGACCAATCAAGATAGTAAGGAGAGCTCACTAAATGCTAAATTAG 540
QY 438 GCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTGAGACAGCGGGAGG 497
DB 541 GCAAAACAGGAGGTAAAGAAATAGCAATCATCTATTGCTGAGACAGCGAGGAGG 600
QY 498 ACAGGATCGGATATTAACCCAGGCACTTCGAGCGGCAACGGCAACCCCTTTGGGTCC 557
DB 601 ACAATGATCAGGATATTAACCCAGGCACTTCAGCCAGCAATGGCTACCTCTTTGGGTCC 660
QY 558 CTTCCCTTTGATGGGCGCTCTGTTTTCATCTATTTCATCTATAAATCTTGCACCT 616
DB 661 CTTCCGTTTGTATGGGAGCTCTGTTTTCATCTATTTCATCTATAAATCTTGCACCT 719
RESULT 30
US-10-104-047-1148
; Sequence 1148, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1148
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-1148
Query Match 73.4%; Score 466; DB 17; Length 2349;
Best Local Similarity 86.4%; Pred. No. 2.5e-134;
Matches 541; Conservative 0; Mismatches 75; Indels 10; Gaps 2;
QY 19 CTTGTTAAGTTTGTCTCTTCAGAACTCAAACTGTAAACTACAAATTTGTTCTTCAAATG 78
DB 1377 CTTCTTCAGTTTGTCTCTTCAGAACTCAAACTGTAAACTACAAATTTGTTCTTCAAATG 1436
QY 79 GAGCACCAGATGGAGTCCATGACTAAGATCCAGTGGACCCCTGGACCGGCTGTAGC 138
DB 1437 GAGCCCCAGATGGAGTCCATGACTAAGATCTACTGCGGACCTCTGGACCGGCTGTAGC 1496
QY 139 CCATGCTCCGATGTTAATGACATTAAGAGCAACCCCTCCGAGGAAATCTCAACTGCACAA 198
DB 1497 CCA-GCTCCCCATTAATGACATTAAGAGCAACCCCTCCGAGGAAATCTCAACTGCATGA 1555
QY 199 CCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGTATACAGCAACCTCCCCA 258
DB 1556 CACCTACTACGCCCAATTCAGCGGGAAGCAGTTAGAGTGTGCTTGGCAACCTCCCCA 1615
QY 259 ACAGCACTGGGTTTCTGTTGAGAGGGGGAGCTGAGACAGGACTAGCTGGATTTCC 318
DB 1616 ACAGCACTGGGTTTCTGTTGAGAGGGGGAGCTGAGACAGGACTAGCTGGATTTCC 1675
QY 319 TAGGCCAAGAGAACTCCCTAAGCCTAGCTGGGAAGGTGACTGATCCACCTCTAAACAT 378
DB 1676 TAGGCCAAGTAAGAACTCCCTAAGCCTAGCTGGGAAGGTGACTACACCCACCTTTAACAC 1735

QY 379 GGGGCTTGCAACTTAGCTCACAACCGACCAATC-----AGAGAGCTCACTAAATG 429
DB 1736 TAGGCTTGCAACTTAGCTCACAACCGACCAATCAGGTAGTAAGAGAGAGCTTGGTAAATG 1795
QY 430 CTAATTAGGCAAAATAGGAGGTAAAGAAATAGCAATCATCTATTGCTGAGAGCACAG 489
DB 1796 CTAATTAGGCAAAACAGGAGGTAAAGAAATAGCAGTCTATTCGCTTGACGACAA 1855
QY 490 CGGGGGGACAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCT 549
DB 1856 GGGGCGGGAACAATGATCAGGATATAAATCAGGCAATCAAGCAGCAATGGCTACCCACT 1915
QY 550 TTGGGTCCCTCCCTTTGATGGGCGCTCTGTTTTCATCTATTTCACCTCTATTAAATCT 609
DB 1916 TTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTTCATCTATTAAATCTTTCACTACC 1975
QY 610 TGCAACTGAAAAAAGAAAAAAGAAAAA 635
DB 1976 AAAAAAAGAAAAAAGAAAAA 2001
RESULT 31
US-10-204-887-10
; Sequence 10, Application US/10204887
; Publication No. US20030124569A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHEN, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; TITLE OF INVENTION: SECRETORY MOLECULES
; FILE REFERENCE: PT-1134 PCT
; CURRENT APPLICATION NUMBER: US/10/204,887
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/185,215; 60/185,216; 60/185,232; 60/205,323; 60/205,287; 60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-05-16; 2000-05-17; 2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1564
; TYPE: DNA


```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030124569A1 LI:813218.1:2000FEB01
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 307, 317-318, 335, 918, 921, 943, 1011
; OTHER INFORMATION: a, t, c, g, or other
US-10-204-887-10

Query Match 70.7%; Score 449; DB 15; Length 1564;
Best Local Similarity 90.9%; Pred. No. 4.2e-129;
Matches 490; Conservative 0; Mismatches 45; Indels 4; Gaps 1;

QY 41 GAATCAAACTGTAAACTACAAATTTGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGA 100
Db 1022 GAATTGAACTGTAAACTACAAATGTTTCATCAAAATGGAGGCCCAAGATGCATCCATGA 1081

QY 101 CTAAAGATCCACCTGGACCCCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAAATGACA 160
Db 1082 CTAAGATCCACCTAGACCCCGGACCGGTCTCCAGCCCATGCTCTGGTGTAAATGACA 1141

QY 161 TTGAAGSCACCCCTCCGAGGAAATCTCAATGCACAAACCCCTACTATGCCCCCAATTGAG 220
Db 1142 TCGAAGSCACCCCTCCCAAGGAAATCTCAGCTGCACAAACCCCTCTATGCCCCCAGTTGAG 1201

QY 221 CGGAAGCAGTTAGAGCGGTTCATCAGCAACCTCCCAACAGCACTTGGGTTTTCCTGTT 280
Db 1202 CAGGAAGCAGTTAGAGCAGTTCATCGGCCAACCTCCCAATAGCACTTGGGTTTTCCTGTT 1261

QY 281 GAGAGGGGAGCTGAGAGACAGGACTAGCTGGATTTCCTAGGCCCAACGAAGAATCCCTAA 340
Db 1262 GAGAGTGGGAGCTAG---AGGACTAGCTGGATTTCCTAGGCCGAGTAAAGATCCCTAA 1317

QY 341 GCCTAGCTGGGAAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACA 400
Db 1318 GCCTAGCTGGGAAGGTAACTACATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACA 1377

QY 401 CCGACCAATCAGAGCTCCTAATAATGCTAATAGGCAAAATAGGAGTAAAGAAAT 460
Db 1378 CCAACCAATCAGAGCTCCTAATAATGCTAATAGGCAAAATAGGAGTAAAGAAAT 1437

QY 461 AGCCAATCATCTATTGCTGAGACACAGCGGAGGAGGACAGGATCGGATATAAACCCTA 520
Db 1438 AGCCAATCATCTATTGCTGAGACACAGCGGAGGAGGACAGGATCAGATATAAATCCA 1497

QY 521 GGCATTCCAGCGGCAACGGCAACCCCTTTGGTCCCTCCCTTTGATGGGCGCTCT 579
Db 1498 GGCATTCCAGCGCAATGGCAACCCCTTTGGTCCCTCTTGTATGGAGCTCTAT 1556

RESULT 32
US-10-220-120-17
; Sequence 17, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
```

```

; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy P.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.
; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKREHER, Theresa K.
; APPLICANT: DAFFO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHER, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 PCT
; CURRENT APPLICATION NUMBER: US/10/220,120
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774
; 60/184,893; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
; 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
; 60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;
; 2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
; 2000-05-17; 2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 1393
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LI:351965.1:2000FEB01
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 120, 765
; OTHER INFORMATION: a, t, c, g, or other
US-10-220-120-17

Query Match 70.3%; Score 446.2; DB 17; Length 1393;
Best Local Similarity 90.3%; Pred. No. 3e-128;
Matches 502; Conservative 0; Mismatches 44; Indels 10; Gaps 2;

QY 47 AAACCTGTAAAACTACAAATTTGTTCTTCAAAATGGAGCACCAGATGGAGTCCATGACTAAGA 106
Db 2 AAGCTGTAAAACTACTAATTTGTTCTTCAAAAGAGAGCCCAAGATGCAGTCCATGACTAAGA 61

QY 107 TCACCGGTGACCCCTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTGAAG 166
Db 62 TCTACCACAGACCCCTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGATATCGANA 121

QY 167 GCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGAA 226
Db 122 GCACCTCC-CCCAAGGAAATTTCACTGCACAAACCCCTACTATACACCCCAATTCAGCAGAA 180

QY 227 GCAGTTAGAGCGGTTCATCAGCCAAACCTCCCAACAGACACTTTGGGTTTTCTGTTGAGAGG 286
Db 181 GCAGTTAGAGCGGTGCTCAGCCAAACCTCCCAACAGACACTTTGGGTTTTCTGTTGAGAGC 240

QY 287 GGGGACTGAGAGACAGGACTAGCTGGATTCTTAGGCCAAGAAATCCCTAAGCCTAG 346
```

Db	241	GGGGA	CTGAGACAGAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCTAAGCCTAG	300
Qy	347	CTGGGAAGGTGACTGCAATCCACTTAACACATGGGGCTTGCAACTTAGCTCACACCCGAC	406	
Db	301	CTGGGAAGGTGACCGGCTACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCCAAC	360	
Qy	407	CAATC-----AGAGAGCTCACTAAATGCTAATTAAGGCAGGAGTAAAGAAATAG	457	
Db	361	CAATCAGGTAGTAAAGAGAGCTCACTAAATGCTAATTAAGGCAGGAGTAAAGAAATAG	420	
Qy	458	AATAGCCCAATCATCTATTTCCTGAGAGCACAGCGGAGGAGCAAGGATCGGATATAAAC	517	
Db	421	AATAGCCCAATCATCTATCGCTGAGAGCACAGCAGGAGGAGCAATGATCCGGATATAAAC	480	
Qy	518	CCAGGCAATTCGAGCCGCGCAACGCCCTTTGGTCCCTCTCTCTTTGATGGGCGCT	577	
Db	481	CCAAGCAATTCGAGCCGCGCAACGGCTGCGCCCTTTGTGTCCCTCTCTCTTTGATGGGAGCT	540	
Qy	578	CTGTTTTCACCTCTATT	593	
Db	541	CTGTTTTCACCTCTATT	556	
RESULT 33				
US-10-198-846-9936				
; Sequence 9936, Application US/10198846				
; Publication No. US2003099974A1				
; GENERAL INFORMATION:				
; APPLICANT: Lillie, James				
; APPLICANT: Xu, Yongyao				
; APPLICANT: Wang, Youzhen				
; APPLICANT: Steinmann, Kathleen				
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS				
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND				
; TITLE OF INVENTION: THERAPY OF BREAST CANCER				
; FILE REFERENCE: MRI-049				
; CURRENT APPLICATION NUMBER: US/10/198,846				
; CURRENT FILING DATE: 2002-07-18				
; PRIOR APPLICATION NUMBER: 60/306,220				
; PRIOR FILING DATE: 2001-07-18				
; NUMBER OF SEQ ID NOS: 14084				
; SOFTWARE: FastSEQ for Windows Version 4.0				
; SEQ ID NO 9936				
; LENGTH: 7974				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-10-198-846-9936				
Query Match 68.4%; Score 434.4; DB 14; Length 7974;				
Best Local Similarity 87.7%; Pred. No. 3.3e-124;				
Matches 537; Conservative 0; Mismatches 56; Indels 19; Gaps 5;				
Qy	1	CCCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCCAGAAATCAAACTGTAAACTA	60	
Db	6409	CCCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCCAGAAATCAAACTGTAAACTA	6468	
Qy	61	CAAAATGTTCTTCAAAATGGAGCACAGATGAGTCCATGAC-TAAGATCCACCGTGGACC	119	
Db	6469	CAAAATGTTCTTCAAAATGGAGCCCAAGATGAGTCCATGACTTAAGATCTACCATGGACC	6528	
Qy	120	CCTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCCGA	179	
Db	6529	CCTGGACCGGCTCTAGCCCATGCTCTGATGTTAATGACATCGAAGACCCCTCCAGA	6588	
Qy	180	GGAAATCTCAACTGCAAAACCCCTACTATGCCCAATTACGGGAGGAGCTAGAGCGG	239	
Db	6589	GGAAATCTCAACTGCAAAACCCCTACTATGCCCTGATTACGAGGAGTAGTAGTGG	6648	
Qy	240	TCATCAGCCAACTCCCAACAGACCTTGGTTTTCTGTTGAGAGGGGGAGCTGAGAGA	299	
Db	6649	TTGTGCGCCAACTCTCCCAACAGACCTTGGTTTTCTGTTGAGAGGGGGAGCTGAGAGA	6708	
Qy	300	CAGGACTAGCTGGATTTCTTA-----GGCCAAAGAAATCCCTAAGCCTAGCTGGGA	352	

Db	6709	CAGGACTAGCTGGATTTCCCTAGGCGCAGCTTAAAGAAATCACCACTAAGCCTAGCTGGGA	6768	
Qy	353	AGGTGACTGAT-CCACCTCTTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATC	411	
Db	6769	AGGTGACCGGCTCCCACTTTAAACACGGGCTTGCAACTTAGCTCACACCCGACCAATC	6828	
Qy	412	-----AGAGAGCTCACTAAATGCTAATTTAGGCAGGAGTAAAGAAATAG	462	
Db	6829	AGGTAGCAAGAGAGCTCGTTAAATGATAATTTAGGCAGGAGTAAAGAAATAG	6888	
Qy	463	CCAATCATC-TATTGCTCTGAGACACAGCGGAGGAGCAAGGATCGGATATAAACCCAG	521	
Db	6889	CCAATCATCTTATCACCTGAGAGCAAAATGGGAGGAGCAATGATCGGATATAAACCCAG	6948	
Qy	522	GCATTCCAGCGGCAACGCCCTTTGGGTCCCTCTCTCTTTGATGGGCGCTCTGT	581	
Db	6949	GCATTCCAGTGGCAATGCTACCTCTTTGGGTCCCTCTCTTTGATGGAGCTCTGT	7008	
Qy	582	TTTCACTCTATT	593	
Db	7009	TTTCACTCTATT	7020	
RESULT 34				
US-10-322-281-718				
; Sequence 718, Application US/10322281				
; Publication No. US20040126762A1				
; GENERAL INFORMATION:				
; APPLICANT: David W. Morris				
; APPLICANT: Marc S. Malandro				
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer				
; FILE REFERENCE: 529452001000				
; CURRENT APPLICATION NUMBER: US/10/322,281				
; CURRENT FILING DATE: 2002-12-17				
; NUMBER OF SEQ ID NOS: 866				
; SOFTWARE: FastSEQ for Windows Version 4.0				
; SEQ ID NO 718				
; LENGTH: 44063				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; NAME/KEY: misc feature				
; LOCATION: (1) - (44063)				
; OTHER INFORMATION: n = A,T,C or G				
US-10-322-281-718				
Query Match 67.3%; Score 427.6; DB 18; Length 44063;				
Best Local Similarity 86.1%; Pred. No. 9.7e-122;				
Matches 556; Conservative 0; Mismatches 54; Indels 36; Gaps 6;				
Qy	1	CCCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCCAGAAATCAAACTGTAAACTA	60	
Db	1026	CCCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCCAGAAATCGAAGCTGTAAACTA	1085	
Qy	61	CAAAATGTTCTTCAAAATGGAGCACAGATGAGTCCATGACTCACTCCCGTGACCC	120	
Db	1086	CAAAATGTTCTTCAAAATGGAGCCCGGATGCACTCCATGACTTAAGATCTTACTGTGGACC-	1144	
Qy	121	CTGGACCGGCTCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCAACCCCTCCGAG	180	
Db	1145	-----CCTGCTAGCCCATGCTCCGATGTTAATGACATCGAAGGCAACCCCTCCGAG	1195	
Qy	181	GAATATCTCAACTGCAAAACCCCTACTATGCCCCCAATTGCGGGAGGAGCTAGAGCGGT	240	
Db	1196	GAATATCAACGCGCATGATCCCTAGTATGCCCAAAATTCAGCAGGAAGC-----AGAGCGGC	1251	
Qy	241	CATCAGCCAACTCCCAACACAGCACTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	300	
Db	1252	CGTCGGCCAACTCTCCCAACACATCACTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGAC	1311	
Qy	301	AGGACTAGCTGGATTTCTCTAGGCCCAACGAAGAAATCCCTAAGCCTAGCTGGG-AAAGGTGAC	359	

Db 1312 AGGACTAACTGGATTTCTTAGGCCAACTAAGAAATCCCAAGCCCTAGCTGGGAAGGTGAC 1371
 Qy 360 TGCAATCAGCTTAAACATGGGGCTTGCACCTTAGCTCACCCGACCAATC----- 411
 Db 1372 TGCAACCCACCTTTAGACATGGGGCTTGTAACTCAGCTCACCCGACCAATCAGGAGTA 1431
 Qy 412 -AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAG-CCAATCA 469
 Db 1432 AAGAGGCTCACTAAATACAAATAGGCTAAGCGAAGGTAAAGAAATAGTCAATCA 1491
 Qy 470 TCTATTGCTCGAGACACAGCGGAGGACAGGATCGGATATAAACCCAGGCAATCGA 529
 Db 1492 TACATTGCTCGAGACACAGGAGGAGCAATATCGGATATAAACCCAGGCAATCGA 1551
 Qy 530 GCGGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGCGCTCTGTTTTCATC 589
 Db 1552 GCAGGAGGCGCAACCCCTTTGGGTCCCTCCCTTTTATGGGAGCTCTG----- 1602
 Qy 590 TATTTCACCTATTAAATCTTGAACCTGAAAAAAGGAAAAA 635
 Db 1603 --TTTCACCTATTAAATCTTGAACCTGAAAAAAGGTTAAA 1646

RESULT 35
 US-10-027-632-289680
 ; Sequence 289680, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 289680
 ; LENGTH: 604
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-289680

Query Match 66.9%; Score 424.6; DB 13; Length 604;
 Best Local Similarity 93.4%; Pred. No. 1.1e-121;
 Matches 453; Conservative 1; Mismatches 30; Indels 1; Gaps 1;
 Qy 110 ACCGTGACCCCTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTGAAGGCA 169
 Db 2 ACCATGACCCCTGGACCGGCTGCTAGCACATGCTCTGATGTTAATGACATCGNAGGCA 61
 Qy 170 CCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCA 229
 Db 62 CCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCAAGAGCA 121
 Qy 230 GTTAGAGCGGTCA-TCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGG 288
 Db 122 GTTAGAGTGGTCACTCGGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGG 181
 Qy 289 GGACTGAGACAGGACTAGTGGATTTCTTAGSCCAACGAAGATCCCTAAGCCTAGCT 348

Db 182 GGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCAACTAAGAAATCCCTAAGCCTAGGT 241
 Qy 349 GGGAGGTTGACTGCATCCACCTCTAAACATGGGGCTTGCACCTTAGCTCACACCCGACCA 408
 Db 242 GGGAGGTTGACTGCATCCACCTTTAAACAGGGGGCTTGCACCTTAGCTCACGCGCGACCA 301
 Qy 409 ATCAGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATC 468
 Db 302 ATAAGAGAGCTTACTAAATGCTAATTAGGCAAAACAGGAGTAAAGAAATAGCCAATC 361
 Qy 469 ATCTATTGCTGAGACACAGCGGAGGACAGGATCGGATATAAACCCAGGCAATCG 528
 Db 362 ATCTATTGCTGAGACACAGTGGGAGGGAATGATCTGGATATAAACCCAGGCAATCG 421
 Qy 529 AGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGCGCTCTGTTTTCACCT 588
 Db 422 AGCCAGCAATGCTAGTCCCTCTATGATCCCTCCCTTTGATGGGAGCTCTGTTTTCACCT 481
 Qy 589 CTATT 593
 Db 482 CTATT 486

RESULT 36
 US-10-027-632-289680
 ; Sequence 289680, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 289680
 ; LENGTH: 604
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-289680

Query Match 66.9%; Score 424.6; DB 17; Length 604;
 Best Local Similarity 93.4%; Pred. No. 1.1e-121;
 Matches 453; Conservative 1; Mismatches 30; Indels 1; Gaps 1;
 Qy 110 ACCGTGACCCCTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTGAAGGCA 169
 Db 2 ACCATGACCCCTGGACCGGCTGCTAGCACATGCTCTGATGTTAATGACATCGAAGGCA 61
 Qy 170 CCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCA 229
 Db 62 CCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCAAGAGCA 121
 Qy 230 GTTAGAGCGGTCA-TCAGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGG 288
 Db 122 GTTAGAGTGGTCACTCGGCCAACTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGG 181

```

QY 289 GGAAGGTGAGTGGATTTCTAGGCCAACGAGAGATCCCTAAGCTAGCT 348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 182 GGAAGGTGAGTGGATTTCTAGGCCAACGAGAGATCCCTAAGCTAGCT 241
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 349 GGAAGGTGAGTGGATTTCTAGGCCAACGAGAGATCCCTAAGCTAGCT 408
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 242 GGAAGGTGAGTGGATTTCTAGGCCAACGAGAGATCCCTAAGCTAGCT 301
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 409 ATCAGAGCTCACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGC 468
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 302 ATAAGAGCTTACTAAATGCTAATAGGCAAAATAGGAGTAAAGAAATAGC 361
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 469 ATCTATTGCTGAGACACAGCGGAGGACAGGATCGGATATAAACCCAGGATCG 528
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 362 ATCTATTGCTGAGACACAGTGGAGGAGGAAATGATCGGATATAAACCCAGGATCG 421
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 529 AGCCGCAACGCGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTCTTTCACT 588
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 482 CTATT 486
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 37
US-10-741-600-17699/c
; Sequence 17699, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17699
; LENGTH: 50353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17699

Query Match 66.2%; Score 420.4; DB 19; Length 50353;
Best Local Similarity 83.9%; Pred. No. 1.8e-119;
Matches 535; Conservative 2; Mismatches 88; Indels 13; Gaps 5;

QY 11 TTAACCTCTGTTAAAGTTTGTCTCTTCCAGAACTCAAACTGTAAACTACAAATGTTC 70
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 30181 TTAACCTCTGTTAAAGTTTGTCTCTTCCAGAACTCAAACTGTAAACTACAAATGTTC 30122
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 71 TTCAATGGAGCCACGATGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGC 130
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 30121 TTCAATGGAGCCACGATGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGC 30062
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 131 CTGCTAGCCCATGCTCGATGTTAATGACATTTGAAGCACCCTCCCGAGGAATCTCAA 190
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 30061 CTGCTAGCCCATGCTCGATGTTAATGACATTTGAAGCACCCTCCCGAGGAATCTCAA 30002
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 CTGCACAAACCTTACTATGCCCCCAATTCAGCGGAAGCAGTTAGAGCGGTCACTACGCCAA 250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 30001 CTGCATGACCCCTTACTATGCCCCCAATTCAGCGGAAGCAGTTAGAGCGGTCACTACGCCAA 29942
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 251 CCTCCCCAACAGCACTTGGTTTTTCTGTTGAGAGGGGGACTGAGAGACAGACTAGCT 310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 29941 CCTCCCCAACAGCACTTGGTTTTTCTGTTGAGAGGGGGACTGAGAGACAGACTAGCT 29882
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 311 GGATTTCTAGGCCAACGAGAGATCCCTAAGCTAGCT-GGAGAGTGAAGTGCATCCACC 369
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 29881 GGATTTCTAGGCCAACGAGAGATCCCTAAGCTAGCT-GGAGAGTGAAGTGCATCCACC 29822
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 370 TCTAAACATGGGGCTTCAACTTAGCTTCACACCCGACCAATC-----AGAGAGCTC 420

```

```

Db 29821 TTTAAACAAGGGCTAGTAACCTCAGCTCACACCTGACCAATCACATAGTAAGAGAGCTC 29762
QY 421 ACTAAATGCTAATAGGCAAAAA-TAGGAGGTAAAGAAAT-AGCCAATCATCTATTGCC 478
Db 29761 ACTAAATATACCAATTAGGCTTAAAGCAGGAGGTAAAGAAATAATCAATAATCTATCACC 29702
QY 479 TGAGAGCACAGCGGAGGACAAAGGATCGGGATATAAA-CCGAGGCAATTCGAGCCGCAA 537
Db 29701 TGAGAGCACAGGAGGAGGACATGATCGGATATAAACCCAGGCAATTCAGCCAGCAG 29642
QY 538 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGCGCTCTGTTTCACTCTATTTCAC 597
Db 29641 TGGCAATTCCTTTGGGCGCCTCCCATTTATATGTTAGCTCTGTTTCACTCTATAAT 29582
QY 598 TCTATTAAATCTTGCACTGAAAAAAGAAAAAAGAAAAA 635
Db 29581 CTTGACGCTGCCAAAAAAGAAAAAAGAAAAAAGAAAA 29544

RESULT 38
US-10-466-531-47
; Sequence 47, Application US/10466531
; Publication No. US20040166500A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, Scott R.
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOUR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy F.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAFPO, Abel;
; APPLICANT: MAWANA, Rakesh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN JR., Edward H.;
; APPLICANT: PERALTA, Careyna H.; DAVID, Marie H.;
; APPLICANT: LEWIS, Samancha A.
; TITLE OF INVENTION: SECRETORY MOLECULES
; FILE REFERENCE: PT-1216 USN
; CURRENT APPLICATION NUMBER: US/10/466,531
; PCT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: PCT/US02/01340
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/262,599
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,131
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,164
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,864
; PRIOR FILING DATE: 2001-01-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PERL Program
; SEQ ID NO 47
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: LI:1147914.1:2001JAN12
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 484
; OTHER INFORMATION: a, t, c, g, or other

```



```
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1175
; LENGTH: 23855
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-355-1175

Query Match      61.2%; Score 388.4; DB 17; Length 23855;
Best Local Similarity 83.1%; Pred. No. 1.3e-109;
Matches 552; Conservative 0; Mismatches 71; Indels 41; Gaps 8;

QY      1  CCCTGTATCTTTAACTCTCTTGAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTTA 60
Db      11414 CCCTGTATCTTTAACTCTCTTGAAGTTTGTCTTCCAGAAATCGAAGCTGTAAACTG 11355

QY      61  CAAATTTGTTCTTCAATGAGCACCAGATGAGTCAATGACTTAAGATCAACCGTGGACCC 120
Db      11354 CAAACAGTTCTTCAATGAGCACCAGATGAGTCAATGACTTAAGATCAACCGGACCC 11295

QY      121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGCACCCCTCCCGAG 180
Db      11294 CTGGACCGGCTGTAGCCCATGCTCTGATGTTAATGACATCGAAGCCACCCCTTCCAAG 11235

QY      181 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTTCAGCGGGAAGCAGTT-AGAGGG 239
Db      11234 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTTCAGCAGGAAGCAGTTAAGATGG 11175

QY      240 TCATCAGCCCAACCTCCCAACAGCACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGA 299
Db      11174 TTGTGGCCCAACCTTGCCCAACAGCACTTGGGCTTTCTGTTGA-CGGGGGGACTGAGAGA 11116

QY      300 CAGGACTAGCTGATTTCTTAGGCCCAACGAAGAAATCCCTAAGCCTAGCTGGG-AGGTGA 358
Db      11115 CAGGACTAGCTGATTTCTTAGGCCGATTAAGAAATTCCTAAGCCTAGCTGGGAAAGGTGA 11056

QY      359 CTGCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 411
Db      11055 CCGCACCTACCTTTAAACACAGGGCTTGTAATCTAGCTCACACCCCAACCAATCAGGTAGT 10996

QY      412 --AGAGAGCTCACTAAATGCTAAATTAGGCAAAATAGGAGGTAAAGAAATAG-CAAATC 468
Db      10995 AAAGAGGGCTCGCTAAATACAAATTAGGCTAGAGCAGAAGGTAAAGACACAGTCAAAATC 10936

QY      469 ATCTATTGCTGAGACACAGCGGAGGACAGAGGATCGGATATAAC-CCAGGCATTC 527
Db      10935 ATATATCGCTGAGGACACAGGGGAGGAGCAATGATGGGATATAAATCTCCAGGCATTC 10876

QY      528 GAGCCGGCA-----ACGGCAACCCCTTTGGGTCCCTCCCTTTGTAT 570
Db      10875 GAGCCGGAGTGGGCAACCTCTTTGAGGGCAACCCCTTTGGGTCCCTCCCAITGTAT 10816

QY      571 GGGCGCTCTGTTTCTACTATTTCATCTATTAAATCTTGCAACTGAAAAAATAAATA 630
Db      10815 GGGAGCTCTGT-----TTTCACTCTATTAAATCTTGCAACTTGCAAAATAATAATA 10765

QY      631 AAAA 634
Db      10765 ATAA 10762
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:52:54 ; Search time 1974.67 Seconds
(without alignments)
12240.400 Million cell updates/sec

Title: US-09-319-156B-6
Perfect score: 635
Sequence: 1 cccgtgatcttttaacctct.....tgaaaaaaaaaaaaaa 635

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST.*
- 1: gb_est1.*
 - 2: gb_est2.*
 - 3: gb_est3.*
 - 4: gb_est4.*
 - 5: gb_est5.*
 - 6: gb_est6.*
 - 7: gb_est7.*
 - 8: gb_est8.*
 - 9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498.8	78.6	1160	3 BC030968	BC030968 Homo sapi
2	492.8	77.6	2748	3 CR605851	CR605851 full-leng
3	488.8	77.0	2749	3 CR617248	CR617248 full-leng
4	488	76.9	2500	3 CR622175	CR622175 full-leng
5	485.6	76.5	771	4 BI087886	BI087886 602852690
6	485	76.4	2748	3 CR625046	CR625046 full-leng
7	470.6	74.1	719	7 CN272394	CN272394 170006000
8	468.6	73.8	1071	5 BX365066	BX365066 BX365066
9	466.6	73.5	758	5 BX357208	BX357208 BX357208
10	464.6	73.2	2716	3 CR613169	CR613169 full-leng
11	454.2	71.5	998	5 BX337769	BX337769 BX337769
12	453.6	71.4	1058	5 BX378303	BX378303 BX378303
13	450.8	71.0	689	9 AG121669	AG121669 Pan.trog
14	437	68.8	494	1 AA781423	AA781423 aj26c03.s
15	435	68.5	1019	5 BX439636	BX439636 BX439636
16	432.4	68.1	522	2 AW971553	AW971553 EST383642
17	430.2	67.7	870	7 CN645411	CN645411 ILLUMIGEN
18	424	66.8	1500	3 BC026287	BC026287 Homo sapi
19	423	66.6	653	9 AG033781	AG033781 Pan.trog
20	415.8	65.5	490	1 AI598135	AI598135 tn4a10.x
21	397.4	62.6	440	2 BE732673	BE732673 601571305
22	396	62.4	543	4 BI963185	BI963185 ie57c10.x
23	395.2	62.2	966	5 BX380176	BX380176 BX380176
24	394.8	62.2	651	9 AG058970	AG058970 Pan.trog

C 25	393.2	61.9	701	9	AG126669	AG126669 Pan.trog
C 26	393	61.9	609	9	AG066901	AG066901 Pan.trog
C 27	392.6	61.8	443	1	AA837267	AA837267 OD26b10.s
C 28	387.2	61.0	712	8	AQ892947	AQ892947 HS_3131.B
C 29	386.8	60.9	436	1	AI128526	AI128526 qc61h10.x
C 30	385.2	60.7	446	1	AI393478	AI393478 tg45g04.x
C 31	384.6	60.6	641	9	AG036829	AG036829 Pan.trog
C 32	384.2	60.5	485	2	AW511366	AW511366 hd45h03.x
C 33	379.8	59.8	664	5	BX481837	BX481837 DXFzp866M
C 34	375.2	59.1	556	1	AUI58595	AUI58595 AUI58595
C 35	369.6	58.2	865	6	CB231128	CB231128 AGENCOURT
C 36	365.8	57.6	470	1	AI074704	AI074704 ox83d05.s
C 37	364.8	57.4	679	9	AG076758	AG076758 Pan.trog
C 38	364.2	57.4	415	1	AI128496	AI128496 qc61e08.x
C 39	360.6	56.8	777	9	AG030228	AG030228 Pan.trog
C 40	359.8	56.7	425	1	AI570707	AI570707 tm79g09.x
C 41	359.8	56.7	431	1	AA552941	AA552941 nk61a10.s
C 42	359.2	56.6	490	1	AA426511	AA426511 zw02e05.r
C 43	359.2	56.6	619	9	AG133542	AG133542 Pan.trog
C 44	359.2	56.6	674	9	AG091649	AG091649 Pan.trog
C 45	355.4	56.0	446	6	CB069106	CB069106 is13a01.x

ALIGNMENTS

BC030968 1160 bp mRNA linear HTC 19-NOV-2003
Homo sapiens cDNA clone IMAGE:4724433, with apparent retained intron.

ACCESSION BC030968

VERSION BC030968.1 GI:22658419

KEYWORDS HTC.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 1160)
Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prance, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shvachenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, Y.S., Krzywinski, M.I., Skalska, U., Smal, S., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

CONTACT

EMAIL

TISSUE

CLONTECH

cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 41 Row: m Column: 10
 This clone has the following problem: retained intron.

FEATURES

Location/Qualifiers
 1..1160
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4724433"
 /tissue_type="Placenta"
 /clone_lib="NIH_MGC_79"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"

ORIGIN

Query Match 78.6%; Score 498.8; DB 3; Length 1160;
 Best Local Similarity 92.8%; Pred. No. 3.6e-133;
 Matches 524; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 70 CTTCAATGAGCAGCAGATGAGTCCATGACTTAAGATCCAGCTGACCCCTGGACCGG 129
 Db 582 CTACAAATGGAGCCCAAGATGAGTCCAGACTAGATCTACGACAGACCCCTGGACCGG 641
 QY 130 CTTGCTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACCCCTCCGAGGAATCTCA 189
 Db 642 CTTGCTAGCCCATGCTGATGTTAATGACATCAAAAGGACCCCTCCTGAGGAATCTCA 701
 QY 190 ACTGCACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGACGGTCAATCAGCCA 249
 Db 702 GTGTCACACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGACGGTCTCGGCCCA 761
 QY 250 ACCTCCCAACAGCAGCTGGGTTTCTGTTGAGAGGGGGGACTGAGACAGGAGCTAGC 309
 Db 762 ACCTCCCAACAGCAGCTTAGGTTTCTGTTGAGAGTTGGGGACTGAGACAGGAGCTAGC 821
 QY 310 TGGATTTCTAGGCCAACGAAGAATCCCTAAGCTAGCTGGGAAGTGTGATCCACCC 369
 Db 822 TGGATTTCTAGGCTGACTAAGAATCCCTAAGCTAGCTGGGAAGTGTGATCCACCC 881
 QY 370 TCTAAACATGGGCTTGCACACTTAGCTCACCCGACCAATCAGAGAGCTCACTAAATG 429
 Db 882 TTTAAACACGGGGCTTGCACACTTAGCTCACCTGACCAATCAGAGAGCTCACTAAATG 941
 QY 430 CTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGCAG 489
 Db 942 CTAATTTAGGCAAAAGACAGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGCAG 1001
 QY 490 CGGAGGGACAAGGATCGGGATATAAACCCAGGCAATTCGAGCGGCAACGGCAACCCCT 549
 Db 1002 CAGGAGGGACAATGATCGGGATATAAACCCAGTCTTCGAGCGGCAACGGCAACCCCT 1061
 QY 550 TTGGGTCCTCCCTCTGTTGATGGGCGCTCTGTTTTCACCTATTTCACCTATTAAATCT 609
 Db 1062 TTGGGTCCTCCCTCTGTTGATGGGAGCTCTGTTTTCATGCTATTTCACCTATTAAATCT 1121
 QY 610 TGCACACTGAATAAAAAAAAAAAAAA 635
 Db 1122 TGCACACTGAATAAAAAAAAAAAAAA 1147

RESULT 2

CR605851 2748 bp mRNA linear HTC 21-JUL-2004
 LOCUS

DEFINITION

full-length cDNA clone CS0DE012YU24 of Placenta of Homo sapiens (human).
 ACCESSION CR605851
 VERSION CR605851.1 GI:50486658
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2748)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> InvitroGen Corporation 1600 Faraday Avenue

REFERENCE

2 (bases 1 to 2748)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES

Location/Qualifiers
 1..2748
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE012YU24"
 /tissue_type="Placenta"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 77.6%; Score 492.8; DB 3; Length 2748;
 Best Local Similarity 90.9%; Pred. No. 2.4e-131;
 Matches 540; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCCTCTTTAAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 60
 Db 2167 CCCTGTATCTTTAACTCCTCTTTAAAGTTTGTCTCTTCAGAAATCAAACTGTAAACTA 2226
 QY 61 CAAATGTTCTTCAATCGAGCAGCAGATGGAGTCCATGACTTAAGATCCACCGTGACCC 120
 Db 2227 -----CAAATGGAGCCCAAGATGCAGTCCAGACTTAAGATCTACCGCAGACCC 2274
 QY 121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACTTGAAGGCACCCCTCCCGAG 180
 Db 2275 CTGGACCGGCTGTAGCCCATGCTGATGTTAATGACTTGAAGGCACCCCTCCCGAG 2334
 QY 181 GAAATCTCAACTGCAACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGACGGT 240
 Db 2335 GAAATCTCAGCTGCACCACTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGACGGT 2394
 QY 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGAC 300
 Db 2395 CGTGGCCCAACCTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGAGCTGAGAGAC 2454
 QY 301 AGGACTAGCTGATTTCTTAGSCCAACGAAGAATCCCTAAGCTAGCTGGGAAGGTGACT 360
 Db 2455 AGGACTAGCTGATTTCTTAGGCTGACTTAAGAAATCCCTAAGCTAGCTGGGAAGGTGACT 2514
 QY 361 GCATCCACCTCTAAACATGGGCTTTGCACTTAGTCTCACCCGACCAATCAGAGAGCTC 420
 Db 2515 ACATCCACCTTTAAACACGGGCTTTGCACTTAGTCTCACCCGACCAATCAGAGAGCTC 2574
 QY 421 ACTAAATGCTTAATTTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCCTG 480
 Db 2575 ACTAAATGCTTAATTTAGGCAAAAGACAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCCTG 2634

```

QY 481 AGAGCACAGCGGAGGACAAAGATCGGATATATAAAACCCAGGATTCGAGCCCGCAACGG 540
    |||
Db 2635 AGAGCACAGGAGGACAAATGATCGGATATATAAAACCCAGGCTTCGAGCCCGCAACGG 2694

QY 541 CAACCCCTTTGGGTCCCTCCCTCTTGTATGGCGCTCTGTTTCACTCTATTT 594
    |||
Db 2695 CAACCCCTTTGGGTCCCTCCCTCTTGTATGGGAGCTCTGTTTCACTCTATTT 2748

RESULT 3
LOCUS CR617248
DEFINITION full-length cDNA clone CS0D1022XJ18 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR617248
VERSION 1 GI:50498055
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2749)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 2749)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1. .2749
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1022XJ18"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 77.0%; Score 488.8; DB 3; Length 2749;
Best Local Similarity 90.8%; Pred. No. 3.4e-130;
Matches 536; Conservative 0; Mismatches 42; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTTTGTTAAAGTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
Db 2172 CCCTGTATCTTTAACTCTTTGTTAACTTTGTCTCTCCAGAAATCAAACTGTAAACTA 2231

QY 61 CAAATTGTTCTTAAATGGAGCACCATGATGGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 2232 -----CAAATGGAGCCCAAGATGTCAGTCCAAGACTAAGATCTACCGCAGACCC 2279

QY 121 CTGGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTTGAAGCACCCTCCCGAG 180
Db 2280 CTGGACCGGCTGTAGCCCATGCTGTGTTTAATGACATTTGAAGCACCCTCCCTGAG 2339

QY 181 GAAATCTCAACTGCACACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240
Db 2340 GAAATCTCAGTGCACACACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 2399

QY 241 CATCAGGCAACCTCCCCAAGACACTTGGGTTTTCTGTTGAGAGGGGGGACTGAGAGAC 300
Db 2400 CGTCGGCAACCTCCCCAAGACACTTAGGTTTTCTGTTGAGATGGGGGACTGAGAGAC 2459

```

```

QY 301 AGGACTAGCTGGATTTCCTAGGCCAACGAAGAAATCCCTAAGCCTAGCTGGAAAGTGACT 360
    |||
Db 2460 AGGACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGAAAGTGACT 2519

QY 361 GCATCCACCTCTAAACATGGGCTTCGAACTTAGCTCACACCGGACCAATCACAGAGCTC 420
    |||
Db 2520 ACATCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACACCTTGACCAATCACAGAGCTC 2579

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAGGAAATAGCCAAATCATCTATTGCTG 480
    |||
Db 2580 ACTAAATGCTAATTAGGCAAAAGACAGAGAGGTAAAGAAATAGCCAATCATCTATTGCTG 2639

QY 481 AGAGCACAGCGGAGGACAAAGATCGGATATATAAAACCCAGGATTCGAGCCCGCAACGG 540
    |||
Db 2640 AGAGCACAGCAGGAGGACAAATGATCGGATATATAAAACCCAGGCTTCGAGCCCGCAACGG 2699

QY 541 CAACCCCTTTGGGTCCCTCCCTCTTGTATGGCGCTCTGTTTCACTCT 590
    |||
Db 2700 CAACCCCTTTGGGTCCCTCCCTCTTGTATGGGAGCTCTGTTTCACTCT 2749

RESULT 4
LOCUS CR622175
DEFINITION full-length cDNA clone CS0D1051YMI3 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR622175
VERSION 1 GI:50502982
KEYWORDS HTC; CNSLT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2500)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 2500)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source Location/Qualifiers
1. .2500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YMI3"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 76.9%; Score 488; DB 3; Length 2500;
Best Local Similarity 91.1%; Pred. No. 5.7e-130;
Matches 534; Conservative 0; Mismatches 40; Indels 12; Gaps 1;

QY 1 CCCTGTATCTTTAACTCTCTTGTAGTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
Db 1925 CCCTGTATCTTTAACTCTCTTGTAGTTGTCTCTCTCCAGAAATCGAAGCTGTAAACTA 1984

QY 61 CAAATTGTTCTTCAATGGAGCACCATGAGTCCATGACTAAGATCCACCGTGACCC 120
Db 1985 -----CAAATGGAGCCCAAGATGTCAGTCCAAGACTAAGATCTACCGCAGACCC 2032

```

FEATURES	source
Location/Qualifiers	
1. 771	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="IMAGE:4993894"	
/cell_line="MGC36"	
/lab_host="DH10B"	
/clone_lib="NIH MGC 10"	
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: Nori; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 kb. Library prepared by Life Technologies."	

RESULT 6	
CR625046	
LOCUS	2748 bp mRNA linear HTC 21-JUL-2000
DEFINITION	full-length cDNA clone CS0DIO44YK06 of Placenta Cot 25-normalized of Homo sapiens (human) .
ACCESSION	CR625046
VERSION	CR625046.1 GI:50505853
KEYWORDS	HTC; CNSLT CDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1. (bases 1 to 2748)
TITLE	Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL	Full-length cDNA libraries and normalization
REMARK	Unpublished
	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue
REFERENCE	2. (bases 1 to 2748)
AUTHORS	Genoscope.
TITLE	Direct Submission

LOCUS	BX365066	1071 bp	mRNA	linear	EST 08-APR-2004
DEFINITION	BX365066 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1046YA18 3-PRIME, mRNA sequence.				
ACCESSION	BX365066				
VERSION	BX365066.2	GI:46304105			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On May 5, 2003 this sequence version replaced gi:30374869.				
RESULT 9	BX357208/c				
LOCUS	BX357208	758 bp	mRNA	linear	EST 08-APR-2004
DEFINITION	BX357208 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1022YJ18 3-PRIME, mRNA sequence.				
ACCESSION	BX357208				
VERSION	BX357208.2	GI:46305595			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On May 5, 2003 this sequence version replaced gi:30376125.				
FEATURES	Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r				
ORIGIN	For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CS1A10122C10NP1&c=4215.r.				
FEATURES	Location/Qualifiers				
source	1. .758 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D1022YJ18" /tissue_type="PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."				
Query Match	73.5%	Score 466.6;	DB 5;	Length 758;	
Best Local Similarity	88.8%	Pred. No. 7.1e-124;	Indels 12;	Gaps 1;	
Matches	509;	Conservative	9;	Mismatches	434;
QY	1	CCCTGTATCTTTAACTCTTGTAACTTTGTCTCTTCGAGATCAAACTGTAAACTA	60		
Db	571	CCCTGTATCTTTAACTCTTGTAACTTTGTCTCTTCGAGATCAAACTGTAAACTA	512		
QY	61	CAAAATTGTTCTTCAAAATGGAGCACCAGATGGATGCTTAAGATCCACCTGGACCC	120		
Db	511	-----CAAATGGAGCCCAAGATGCTTAAGATCAAAAGCACCCTCTCTGAG	464		
QY	121	CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATGAAGGCACCCCTCCGAG	180		
Db	463	CTGGACCGGCTGCTAGCCCATGCTCCGATGTTTAATGACATGAAGGCACCCCTCTGAG	404		
QY	181	GAATCTCACTGACACACCCCTACTATGCCCAATTCAGGGGAGCAGTTAGACGGT	240		
Db	403	GAATCTCAGTGCACACCTCTACTACGCCCAATTCAGAGGAGCAGTTAGACGGT	344		
QY	241	CATCAGCCAACTCTCCCAACAGCACTTTGGTTTTCCTGTTGAGAGGGGAGCTGAGAG	300		
Db	343	CTCGGCCAACTCTCCCAACAGCACTTAGGTTTTCCTGTTGAGATGGGGACTGAGAG	284		
QY	301	AGGACTAGTGAATTTCTAGGCCAACGAGAAATTCCTTAGCCCTAGCTGGGAGGTGACT	360		
Db	283	AGGACTAGTGAATTTCTAGGCCGACTAAGAAATTCCTTAGCCCTAGCTGGGAGGTGACC	224		
QY	361	GCATCCACTCTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAATCAGAGGCTC	420		
Db	223	ACATCCACTTTAAACATGGGGCTTGCACTTAGCTCACACCTGACCAATCAGAGGCTC	164		

Db	408	GAAATCTCAGCTGCACAACTCTACTACGCCCAAAATTCAGCAGGAAGCACTTAGAGCGGT	349
Qy	241	CATCAGCCACCTCCCAACAGCACATCTGGGTTCCTGTTTGAAGGGGGGACCTGAGAGAC	300
Db	348	CGTGGGCAACTCCCAACAGCACCTTAGGTCTTCTGTTTGAATGGGGACCTGAGAGAC	289
Qy	301	AGGACTAGCTGGATTTCCCTTAGGCCCAACGAAGAATCCCTAAGCCTTAGCTGGGAAGGTGACT	360
Db	288	AGGACTAGCTGGATTTCCCTAGGCTGACTAAGAATCCCTAAGCCTTAGCTGGGAAGGTGACC	229
Qy	361	GCATCCACCTCTAAACATGGGGCTTGCAAATTAGCTCACACCGACCAATCAGAGAGCTC	420
Db	228	ACATCCACCTTTAAACACGGGGCTTGCACTTAGCTCACACCTGACCAATCAGAGAGCTC	169
Qy	421	ACTAAAAATGCTAATTAGGCGAAAAATAGGAGGTAAAGAAATAGGCAATCATCTATTGGCTG	480
Db	168	ACTAAAAATGCTAATTAGGCGAAAAACAGAGGTAAAGAAATAGGCAATCATCTATTSSCYs	109
Qy	481	AGAGCAGCGGGAGGCAAGGATCGGGATATAAACCCAGGCATTCGAGCCCGCAACGG	540
Db	108	AGAGCAGCAGGAGGGAACAATGATCGGATATAAACCCMMSYCCCGCGCGGCCACGG	49
Qy	541	CAACCCCGTTTGGGTCCCGCTTCTGTATGGG	573
Db	48	CMACCCCCYCTGGGTCCCGCTTCTGTATGGG	16

RESULT 10
LOCUS CR6131169
DEFINITION full-length cDNA clone CSODE013Y120 of Placenta of Homo sapiens (human).
ACCESSION CR6131169
VERSION CR6131169.1 GI:50493976
KEYWORDS HTC; CNSLT CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 2716)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

ORIGIN

Query Match	73.2%	Score	464.6	DB	3	Length	2716
Best Local Similarity	90.9%	Pred. No.	3.6e-123				
Matches	510	Conservative	0	Mismatches	39	Indels	12
Gaps							

QV 1 CCCTGTATCTTTAAACCTCCTTGTGTAAGTTTGTCTCTCCAGAATCAAAACGTGTAACCTA 60

	Db	 CCTGTATCTTTAACTCTCTTGTTAACTTGTCTTCAGAAATCGAAGCTGTAAACTA	2227
2168	Qy	CAAATTGTTCTTCAAAATGGACACCAAGATGGAGTCCATGACTTAAGATCCACCCTGGACCC	120
2228	Db	-----CAAATGGAGCCCAAGTAGCATCCAAGACTAAGATCTACCGCAGACCC	2275
121	Qy	CTGGACCGGCCTGCTAGCCCCATGCTCCGATGTTTAATGACATGAAGACACCCCTCCCGAG	180
2276	Db	CTGGACCGGCCTGCTAGCCCCACGATCTGATGTTTAATGACATCAAAGGCACCCCTCCTGAG	2335
181	Qy	GAAATCTCAACTGCAACAACCCCTACTATATGCCCCAAATTCACGGGGAAGCAGTTAGAGCGGT	240
2336	Db	GAAATCTCAGCTGCAACAACCTCTACTACGCCCCAAATTCACGAGAAGCAGTTAGAGCGGT	2395
241	Qy	CATCAGCGAACCTCCCCACACAGCACACTTGGTTTTCTGTTGAGAGGGGGACACTGAGAGAC	300
2396	Db	CGTCGGCCAACTCCCACACAGCACACTTAGGTTTTCTGTTGAGATGGGGACACTGAGAGAC	2455
301	Qy	AGGACTAGCTGGGATTTCTTAGGCCAAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT	360
2456	Db	AGGACTAGCTGGGATTTCTTAGGCTGACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACC	2515
361	Qy	GCATCCACCTCTAAAACATGGGGCTTGCAACTTAGCTCACCCGACCAATCAGAGAGCTC	420
2516	Db	ACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTC	2575
421	Qy	ACTAAATGCTAATTAGGCCAAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCGTG	480
2576	Db	ACTAAATGCTAATTAGGCCAAAGACGAGAGGTAAAGAAATAGCCAATCATCTATTGCGTG	2635
481	Qy	AGAGCACAGCGGGAGGGAACAAGGATCGGGATAATAAACCCAGGCAATTCGAGCCGGCAACGG	540
2636	Db	AGAGCACAGCAGGAGGGACAATGATCGGGATAATAAACCCAGCTTCGAGCGCGCAACGG	2695
541	Qy	CAACCCCTTTGGGTCCCCCTC	561
2696	Db	CAACCCCTTTGGGTCCCCCTC	2716

RESULT	11
LOCUS	BX337769/c
DEFINITION	BX337769 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODIO51YM13 3-PRIME, mRNA sequence.
ACCESSION	BX337769
VERSION	BX337769.2 GI:46272079
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 998) Li,W.B., Gruber,C.; Jesse,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001) On May 2, 2003 this sequence version replaced gi:30337641.
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

```

FEATURES
  source
    1..998
      /organism="Homo sapiens"
      /mol_type="mRNA"
4215..r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdnar7a=CS0DI05IAG07NP1&c=4215..r

```

FEATURES
source

/db_xref="taxon:9606"
/clone="CS0DI015YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 71.5%; Score 454.2; DB 5; Length 998;
Best Local Similarity 89.1%; Pred. No. 3e-120;
Matches 525; Conservative 2; Mismatches 47; Indels 15; Gaps 3;

QY 1 CCTGTATCTTAAACCTCTTGTAAAGTTTCTCTCCAGATCAAAACCTGTAATACTA 60
DB 579 CCTGTATCTTAAACCTCTTGTAAAGTTTCTCTCCAGATCAAAACCTGTAATACTA 520

QY 61 CAAATTGTTCTTCAATGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 519 -----CAATGGAGCCCAAGATGCGATCCAGACTAAGATCTACCGCAGACCC 472

QY 121 CTGACCGCGCTGCTAGCCATGCTCCGATGTTAATGACATTGAA-GGCACCCCTCCCGA 179
DB 471 CTGGACCGGCTGCTGCGCCACGATCTGATGTTAATGACATCAAAACCGCACCCTCTGA 412

QY 180 GGAATCTCACTGACACACCCCTACTATGCCCCCAATTCAGCGGAGCAGTTAGAGCGG 239
DB 411 GGAATCTCAGCTGACACACCTCTACTAGCGCCCAATTCAGCAGGAAGCAGTTAGAGCGG 352

QY 240 TCATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTCTGTGAGGGGGGAGTCAAGAGA 299
DB 351 TCGTCGGCCAACTCCCAACAGCAGCTTAGGTTTCTCTGTGAGATGGGGAGTCAAGAGA 292

QY 300 CAGGACTAGCTGATTTCTTAGCCCAACGAAGAATCCCTTAAGCTTAGCTGGGAAGTGAC 359
DB 291 CAGGACTAGCTGATTTCTTAGCTGACTAAGAATCCCTTAAGCTTAGCTGGGAAGTGAC 232

QY 360 TGCATCCACTTAAACATGGGGCTTGCACCTTAGCTACACCCGAGCAATCAGAGAGCT 419
DB 231 CACATCCACTTAAACATGGGGCTTGCACCTTAGCTACACCTGAGCAATCAGAGAGCT 172

QY 420 CACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCT 479
DB 171 CACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCT 112

QY 480 GAGACACAGCGGAGGACCAAGGATCGGATATAAACCAGGCAATTCAGCGCGGCAAG 539
DB 111 GAGACACAGCAGGAGGACCAATGATCGGATATAAACCAGGCAATTCAGCGCGGCAAG 52

QY 540 GCAACCCCTTTGGGTCCC--CTCCCTTTGTATGGGCGCTCTGTTTCA 586
DB 51 GCAACCCCTTTGGGTCCCNCCCTTTGTANTGGGAGGTCTGTTTCA 3

RESULT 12
BX378303/c 1058 bp mRNA linear EST 23-APR-2004
LOCUS
DEFINITION
BX378303 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI018YH07 3-PRIME, mRNA sequence.
ACCESSION
BX378303
VERSION
BX378303.2 GI:46557492
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1058)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30439129.
Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DI018CD04NP1&c=4215.r.

FEATURES

Location/Qualifiers
1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI018YH07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 71.4%; Score 453.6; DB 5; Length 1058;
Best Local Similarity 87.0%; Pred. No. 4.5e-120;
Matches 508; Conservative 1; Mismatches 63; Indels 12; Gaps 1;

QY 1 CCTGTATCTTAAACCTCTTGTAAAGTTTCTCTCCAGATCAAAACCTGTAATACTA 60
DB 576 CCTGTATCTTAAACCTCTTGTAAAGTTTCTCTCCAGATCAAAACCTGTAATACTA 517

QY 61 CAAATTGTTCTTCAATGAGCAGCAGATGGAGTCCATGACTAAGATCCACCGTGGACCC 120
DB 516 -----CAATGGAGCCCAAGATGCGATCCAGACTAAGATCTACCGCAGACCC 469

QY 121 CTGACCGCGCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAAGGCACCCCTCCCGAG 180
DB 468 CTGACCGCGCTGCTAGCCCATGCTCCGATGTTAATGACATTGAAAGGCACCCCTCCCGAG 409

QY 181 GAAATCTCAACTGACACACCCCTACTATGCCCAATTCAGCGGAGCAGTTAGAGCGGT 240
DB 408 GAAATCTCAGCTGACACACCTCTACTACGCCCAATTCAGCAGGAAGCAGTTAGAGCGGT 349

QY 241 CATCAGCAACCTCCCAACAGCAGCTTGGGTTTCTCTGTGAGGGGGGAGCTGAGAGAC 300
DB 348 COTCGGCCAACCTCCCAACAGCAGCTTAGGTTTCTCTGTGAGATGGGGAGCTGAGAGAC 289

QY 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAGAATCCCTTAAGCTTAGCTGGGAAGGTGACT 360
DB 288 AGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTTAAGCTTAGCTGGGAAGGTGACT 229

QY 361 GCATCCACCTCTAAACATGGGGCTTGCACTTAGCTACACCCGAGCAATCAGAGAGCTC 420
DB 228 ACATCCACCTTTAAACACGGGCTTTGCACTTAGCTACACCTGAGCTACCAATCAGAGAGCTC 169

QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCTCTG 480
DB 168 ACTAAATGCTAATTAGGCAAAAGCAGAGGTAAAGAAATAGCCCAATCATCTATTGCTCTG 109

QY 481 AGAGCACAGCGGAGGAGCAAGGATCGGGATATAAACCAGGCAATTCAGAGCGGCAAGCG 540
DB 108 AGAGCACAGCAGGAGGAGCAATGATCGGGATATAAACCAGGCAATTCAGAGCGGCAAGCG 49

QY 541 CAAACCCCTTTGGGTCCCCTCCCTTTGTATGGGCGCTCTGTTT 584
DB 48 CAAACCCNNTNNTGGTNNCNCNTGGCGTTTGTGTGGAGTTTGTGTT 5

RESULT 13
AG121669/c
LOCUS

AG121669 689 bp DNA linear GSS 04-NOV-2001


```

DEFINITION Pan troglodytes DNA, clone: PTB-130M15.F, genomic survey sequence.
ACCESSION AG121669
VERSION AG121669.1 GI:16650834
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totohi,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 689)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totohi,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimps@gsc.riken.go.jp URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
Location/Qualifiers
1..689
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-130M15.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 71.0%; Score 450.8; DB 9; Length 689;
Best Local Similarity 88.3%; Pred. No. 2.7e-119;
Matches 515; Conservative 0; Mismatches 58; Indels 10; Gaps 2;

Qy 1 CCCTGTATCTTTAACTCTTGTGTTAGTTTGTCTTTCCAGATCAAACTGTAATACTA 60
Db 666 CCNTGGATCTTTAACTCTTGTGTTAGTTTGTCTTTCCAGATCAAACTGTAATACTA 607
Qy 61 CAATTTGTTCTTCAATGGAGCACCAGATGGAGTCCATGACTAGATCCACCGTGGACCC 120
Db 606 TAATAGTTTCTTCAATGGAGCACCCTTATGTCAGTCCATGACTAGATCTACCATGGACCC 547
Qy 121 CTGACCGGCTCTAGCCCATGCTCCGATGTTTAAATGACATTGAAGCACCCTCCCGAG 180
Db 546 CTGACCGGCTCTAGCCCATGCTCTGATGTTTAAATGACATCGAGCACCCTCTGAG 487
Qy 181 GAAATCTCAATGTCAC-AAACCTTACTATGCCCAATTTCAGCGGGAAGAGTTAGACGG 239
Db 486 GAAATCTCAATGTCACAAACCTTACTTACACTCCAGTTCAGCAGGAAGAGTTAGACGG 427
Qy 240 TCATCAGCACTCTCCCAACAGCACTTGGTGTTCCTGTTGAGAGGGGGAGTACGAGA 299
Db 426 TCGTCGGCACTCTCCCAATGGCACTTGGGTGTTCCTGTTGAGAGGGGGAGTACGAGA 367
Qy 300 CAGGACTAGCTGGATTTCTTAGGCCAACAGAGATCCCTAGCCCTAGCTGGGAAGGTGAC 359
Db 366 CAGGACTATCTGGATTTCTTAGGCCAACATAGATCCCTAGCCCTAGCTGGGAAGGTGAC 307
Qy 360 TGCATCCACCTCTAAACATGGGGTGTGCAACTTAGCTCACACCCGACCAATC----- 411
Db 306 CACATTCACTTTTAAACACAGGGCTTGCACCTTAGCTCACACCCGACCAATCAGGTAGTA 247

```

```

Qy 412 -AGAGAGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCAT 470
Db 246 AAGAGGGCTCACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCAT 187
Qy 471 CTATTGCTTGGAGACACAGCGGAGGACAGGATCGGATATTAACCCAGGATTCGAG 530
Db 186 TTATTGCTTGGAGATACAGCGGAGGACAGGATCGGATATTAACCTATGCAATTCGAG 127
Qy 531 CCGGCAACGGCAACCCCTTTGGGTCCCTCCCTCTTGTATGGG 573
Db 126 CCACCAATGCTACCCCTTTGGGTCCCTCTTGTATGG 84

RESULT 14
LOCUS AA781423/c
DEFINITION aa781423.1 Soares_testis_NHT Homo sapiens cDNA clone 1391428 3',
similar to contains PTR7.tl PTR7 repetitive element ;, mRNA
sequence.
ACCESSION AA781423
VERSION AA781423.1 GI:2840754
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 494)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldino, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 1645 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 475.
FEATURES
Location/Qualifiers
1..494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="1391428"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares_testis_NHT"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [s].
TGTTACCAATCTGAGTGGAGCGGCGCCCAATTTTATTTT 3'".
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldino."
ORIGIN
Query Match 68.8%; Score 437; DB 1; Length 494;
Best Local Similarity 92.7%; Pred. No. 2.5e-115;
Matches 456; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 131 CTGCTAGCCCATGCTCGATGTTAATGACATTAAGGACCCCTCCGAGGAATCTCAA 190
Db 494 CTGCTAGCCCATGCTCGATGTTAATGACATTAAGGACCCCTCCGAGGAATCTCAA 435

```

double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN	Query Match	Score	DB 5	Length	1019
	Best Local Similarity	83.4%	Pred. No. 1.1e-114		
	Matches	494	Conservative	19	Mismatches 66; Indels 13; Gaps 2;
Qy	1	CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA	60		
Db	579	CCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA	520		
Qy	61	CAATGTTCTTCAATGGAGCACAGATGAGTCCATGATTAAGATCCACCGTGGACCC	120		
Db	519	-----CAATGGAGCCCAGATGAGTCCATGATTAAGATCCACCGTGGACCC	472		
Qy	121	CTGACCGGCTGTAGTCCCATCTCGATGTTAAAGATTAAGAGCAGCCCTCCCGAG	180		
Db	471	CTGACCGGCTGTAGTCCCATCTCGATGTTAAAGATTAAGAGCAGCCCTCCCGAG	412		
Qy	181	GAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGAGCAGTTAGAGCGGT	240		
Db	411	GAATCTCAACTGCACAAACCCCTACTATGCCCCAAATTCAGCGGAGCAGTTAGAGCGGT	352		
Qy	241	CATCAGCAACCTCCCCCAACAGCAGCTTTGGGTTTTCTGTGTGAGAGGGGAGCAGT	300		
Db	351	CGTGGCAACCTCCCCCAACAGCAGCTTTAGGTTTTCTGTGTGAGAGGGGAGCAGT	292		
Qy	301	AGGACTAGTGGATTTCTTAGGCCAACAGAAATCCCTAAGCCTAGCTGGGAGGTGACT	360		
Db	291	AGGACTAGTGGATTTCTTAGGCCAACAGAAATCCCTAAGCCTAGCTGGGAGGTGACT	232		
Qy	361	GCATCCACCTTAACATGGGCTTGCAATTTAGCTCACACCCCAACCAATCAGAGAGCTC	420		
Db	231	ACATCCACCTTTAAACAGCGGCTTGCAATTTAGCTCACACCCCAACCAATCAGAGAGCTC	172		
Qy	421	ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTG	480		
Db	171	ACTAAATGCTTAATTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTG	112		
Qy	481	AGAGCAGCGGGAGGACAGGATCGGATATAAACCAGGCAATTCGAGCGGCAACGG	540		
Db	111	AGAGCAGCAGGAGGAGGACATATCGGGATATAAAYCAAAGTCTTCGAGCGGCAACAG	52		
Qy	541	CAACCCCTTTGGGTCCTCCCTTTGTATGGGCTCTGTCTTTCACCTAT 592			
Db	51	AAACCCCTTTT-TTTCYCCCTTTTGTGTTGGGNDTTTTTTTTTNNTTTATTT 1			

RESULT 16
 AW971553/c
 LOCUS
 DEFINITION
 EST383842
 AW971553
 AW971553.1
 GI:8161399
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 522)
 Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
 Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and
 Quackenbush, J.
 Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)
 Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208

Qy	191	CTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGTTCATCAGCAA	250
Db	434	CTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGTTCGTCGGCAA	375
Qy	251	CCTCCCAACAGCAGCTTGGTTTTCTGTGAGAGGGGGAGCTGAGACAGGACTAGCT	310
Db	374	CCTCCCAACAGCAGCTTGGTTTTCTGTGAGATGGGGAGCTGAGACAGGACTAGCT	315
Qy	311	GGATTTCTTAGCCACAGGAATCCCTAAGCTAGCTGGGAAGTGAAGTGCATCCACT	370
Db	314	GGATTTCTTAGCCACAGGAATCCCTAAGCTAGCTGGGAAGTGAAGTGCATCCACT	255
Qy	371	CTAAACATGGGCTTGCACTTAGCTCACACCCGCAACCAATCAGAGAGTCACTAAATGC	430
Db	254	TTAAACAGGGCTTGCACTTAGCTCACACCCGCAACCAATCAGAGAGTCACTAAATGC	195
Qy	431	TAAATTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTGAGAGCAGC	490
Db	194	TAAATTAGGCAAAATAGGAGGTAAGAAATAGCAATCATCTATTGCTGAGAGCAGC	135
Qy	491	GGGAGGACAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACCGCCCTT	550
Db	134	AGGAGGACAGGATCGGATATAAACCCAGGCAATTCGAGCGGCAACCGCCCTT	75
Qy	551	TGGGTCCTCCCTCTTGTATGGGCTCTGTGTTCACTCTATTCACTATTAATCTT	610
Db	74	TGGGTCCTCCCTCTTGTATGGGCTCTGTGTTCACTCTATTCACTATTAATCTT	15
Qy	611	GCAACTGCAAAAAA 624	
Db	14	GCAACTGCAAAAAA 1	

RESULT 15
 BX439636/c
 LOCUS
 DEFINITION
 BX439636 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012VJ24
 3-PRIME, mRNA sequence.
 ACCESSION
 BX439636
 BX439636.2
 GI:47000005
 EST.
 SOURCE
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1019)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 15, 2003 this sequence version replaced gi:30771765.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdnas?cs=CS0DE012D812NP1&c=4215.r.
 Location/Qualifiers
 1. 1019
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE012VJ24"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="vector: pCMVSPORT_6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,

FEATURES
 source

Email: johnq@tigr.org

Plate: 292

Seq primer: Forward

Location/Qualifiers

1. 522
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="WAGE ressequencing, MAGL"
/note="vector: pBluescriptSKm"

FEATURES

Source

ORIGIN

Query Match 68.1%; Score 432.4; DB 2; Length 522;
Best Local Similarity 91.8%; Pred. No. 5.5e-114;
Matches 479; Conservative 0; Mismatches 41; Indels 2; Gaps 2;
Qy 97 ATGACTAAGATCCACCGTGACCCCT-GGACCGGCTGTAGCCCATGCTCCGATGTAA 155
Db 522 AGGTTTAAAGATCTTCGCGGACCCCTAAACCTGGCTTGCTAGCCCATGTTCTGGTGTAA 463
Qy 156 TGACATTTGAAGGACCCCTCCGAGGAA-TCTCAACTGCACACCCCTACTATGCCCA 214
Db 462 TGACATCGAATCTCTCTCTGAGGAAATCTCAACTGCACACCCCTTATTATGCCCA 403
Qy 215 ATTGAGCGGGAAGCAGTTAGAGCGGTCTAGCCCACTCCCAACAGCACTTGGGTTT 274
Db 402 GTTCAGCAGGAGGAGTTAGGTGCTCATCGCCAACTCACCACAGCACTTGGGTTT 343
Qy 275 CTTGTTGAGAGGGGGAGTCTGAGACAGGACTAGCTGGATTTCTAGGCCAACGAGAA 334
Db 342 CTTGTTGAGAGGGGGAGTCTGAGACAGGACTAGCTGGATTTCTAGGCCGATTAAGAA 283
Qy 335 CCCTAAGCTAGTGGGAGGTGCTGATCCACCTCTAACATGGGGCTTGCAACTAG 394
Db 282 CCCTAAGCTAGTGGGAGGTGCTGATCCACCTCTAACATGGGGCTTGCAACTAG 223
Qy 395 CTCACCCGACCAATCAGAGAGTCTCAATAATGCTAAATAGGCAAAATAGAGGTAA 454
Db 222 CTCACCCGACCAATCAGAGAGTCTCAATAATGCTAAATAGGCAAAATAGAGGTAA 163
Qy 455 AGAATAGCCCAATCATCTATTGCTGAGACAGCAGCGGGGACAGGATCGGATATA 514
Db 162 AGAATAGCCCAATCATCTATTGCTGAGACAGCAGTGGGAGGACAGGATTCGAATA 103
Qy 515 AACCCAGCATTCAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGC 574
Db 102 AACCCAGCATTCAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGG 43
Qy 575 GCTGTTTTCACCTCTATTTCACCTCTATTAAATCTTGCAACT 616
Db 42 GCTGTTTTCACCTCTATTTCACCTCTATTAAATCTTGCAACT 1

RESULT 17

CN645411 870 bp mRNA linear EST 13-MAY-2004
LOCUS ILLUMIGEN MQ_23968 Katze_MMSP Macaca mulatta cDNA clone
DEFINITION IBUW:10227 5' similar to Bases 165 to 770 highly similar to human
Unigene Hs.349001, mRNA sequence.
CN645411
CN645411.1 GI:47158854

ACCESSION

VERSION

KEYWORDS

SOURCE

Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.

REFERENCE

AUTHORS Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
TITLE Large-scale Rhesus Macaque cDNA Sequencing
JOURNAL Unpublished (2003)
CONTACT: C. Magness
COMMENT Illumigen Biosciences Inc.

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA

Tel: 2063780400

Fax: 2063780408

Email: cmagness@illumigen.com

Sequenced on 2004.02.24. 676 Q20 bases.

PCR Primers

FORWARD: CCTCTACTAAGGGAACAAA

BACKWARD: CACTATAGGCGAATTGGTA

Insert Length: 870 Std Error: 0.00

Plate: CL000135 row: H column: 02

Seq primer: CCTCTACTAAGGGAACAAA

POLYA=No.

Location/Qualifiers

FEATURES

Source

1. 870
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBUW:10227"
/sex="male"
/cell_type="mononuclear lymphocyte"
/dev_stage="adult"
/lab_host="E. coli SOLB"
/clone_lib="Katze_MMSP"
/notes="Organ: spleen; Vector: Uni-ZAP XR; Site: 1: EcoR I;
Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis
kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold
Cloning kit (Catalog #200450)"
ORIGIN
Query Match 67.7%; Score 430.2; DB 7; Length 870;
Best Local Similarity 83.8%; Pred. No. 2.6e-113;
Matches 527; Conservative 0; Mismatches 88; Indels 14; Gaps 3;
Qy 21 TGTTAAGTTGTTCTCTCCAGAAATCAAACTGTAAACTACAAATGTTCTTCAATGGA 80
Db 166 TTTTAAATTTGTTCTCTCCAGAAATCGAACTGTAAACTACAAATGTTCTTCAATGGA 225
Qy 81 GCACCATGAGGTCCATGACTCAATGATCCACCGTGGACCCCTGGACCGGCTGCTAGCCC 140
Db 226 GCCTCCGATCGATCCATGACTAAGTCTACTGGGACCCCTGGACCGGCTGCTATACC 285
Qy 141 ATGCTCCGATGTTAATGACATTAAGGACACCCCTCCGAGGAAATCTCAATGCACAAAC 200
Db 286 ATGCTCTGATGTTGGTGACATCGAAGGACCCCTCCGAGGAAATCTCAATGCATGACC 345
Qy 201 CCTACTATGCCCAATTCAGGGGAGCAGTGTAGCGGTCATCAGCCCACTCCCAAC 260
Db 346 CCTACTACGCTCAGTTTCAGCAGGACGATGTAGAGCAGTCATCAGCCCACTCCTCAAC 405
Qy 261 AGCACTTGGGTTTTCCTGTTGAGAGGGGGAGCTGAGAGACAGGACTAGCTGGATTTCCTA 320
Db 406 AGCACTTGGATTTCCTGTTGTAAGGGGAGCAGTGTAGAGACAGGACTAGCTGGATTTCCTA 465
Qy 321 GGCCAAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTGATCCACCTCTAAACATGG 380
Db 466 AGCCAACTAAGAATCCCTAAGCCTAGCTGGGAAGGTGACTGATCCATCTTTAAACACGG 525
Qy 381 GGCTTGCAACTTAGCTCAGCCCGGACCAATC-----AGAGAGCTCACTAAATGCT 431
Db 526 GGCTGCAATTTAGCTCAGCCCGGACCAATCAATCAGGTAGTAAAGAGAGCTCCTAAATGTT 585
Qy 432 AATTAGGCAAAATAGGAGGTAAAGAAATAGCCAAATCATC--TATTGCTGAGAGCAGCAG 489
Db 586 AATTAGCAAAACAGGAGGTAAAGAAATAGTCAATCACCTTATCGCTGAGAGCAGCAG 645
Qy 490 CGGAGGAGCAAGGATCGGATATAAACCCAGGCAATTCGAGCCCGGCAACGGCAACCCCT 549
Db 646 GAGGAAGGACAAATGATAGGATATAAACCCAGGCAATTCGAGCAATTCGAGCAATTCGCAACCCCT 705
Qy 550 TT----GGGTCCCTCCCTTTGTATGGGGCTCTGTTTTCACCTCTCTTCTCTCTATTAA 606
Db 706 TTTTGGTCCCTCCCTCCCAATTTTATGGGAGCTCTGGGTTTCTCTATTAAATCTTTCGCAATT 765

```

Qy 607 TCTTCCAACTGAAAAAATATAAATAAATAA 635
Db 766 GCTCAAAAAAACAATATATAAATAAATAA 794

RESULT 18
BC026287 1500 bp mRNA linear HTC 04-MAR-2003
LOCUS Homo sapiens, clone IMAGE:4770655, mRNA.
DEFINITION BC026287
ACCESSION BC026287.1 GI:22382161
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Strausberg, R.
JOURNAL Direct Submission
Submitted (02-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 37 Row: m Column: 12
This clone has the following problem: retained intron.
FEATURES
    source
        1..1500
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4770655"
            /tissue_type="Placenta"
            /clone_lib="NIH MGC 79"
            /lab_host="DH10B"
            /note="Vector: pDNR-LIB"

ORIGIN
Query Match 66.8%; Score 424; DB 3; Length 1500;
Best Local Similarity 85.7%; Pred. No. 1.9e-111;
Matches 557; Conservative 0; Mismatches 75; Indels 18; Gaps 7;

Qy 1 CCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTCCAGATCAAACTGTAAACTA 60
Db 849 CCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTCCAGATCAAACTGTAAACAA 908

Qy 61 CAAATTTCTTCAAAATGGAGCACCAGATCGAGTCCATGATCAAGATCCACCGTGACCC 120
Db 909 CACATCGTTCTTCAAAATGGAGCACCAGATCGAGTCCATGATCAAGATCAACCGGTAC 968

Qy 121 CTGACCGCGCTGTAGCCCATGCTCCGATGTTTAATGACATCAAGACACCCCTCCCGAG 180
Db 969 CTGACCGAGCTGTAGCCCATGCTCCGATGTTTAATGACATCAAGACACCTCTCCCGAG 1028

Qy 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCGGAGCAGTTAGCGGT 240
Db 1029 GAAATCTCAATGCACAAACCCCTACTATGCCCCCAATTAAGCAGGAAGCAGTTGGAGCAGT 1088

```

```

Qy 241 CATCAGCCAACTCCCAACAGCACTTGGGTTTTCCTGTGAGAGGGGGGACTGAGAGAC 300
Db 1089 CGAGGGCCAACTCCCAACAGCACTTGGGTTTTCCTGTGAGAGGGGGTACTGAGAGAC 1148

Qy 301 AGGACTAGCTGGATTTCCTAGGCGCAACGAGAGATCCCTAAGCCCTAGCT-GGGAGGTGAC 359
Db 1149 AGGACTAGCTGGATTTCCTAGGCGCGACTAAACAATTTCTAAGCTTGGTGGGGAAGGTGAC 1208

Qy 360 TGCATCCACTCTAAACATAGGGGC--TTGCAACTTAGCTCACACCCGACCAATC----- 411
Db 1209 TGCACCTACCTTTAAACACAGGACTTTGTAACTCAGCTCACATCTGCGCAATCAGGTAG 1268

Qy 412 ---AGAGAGCTCACTAAATGCTAATPAGGC-AAAATAGGAGGTAAAGAAATAGCCAAT 467
Db 1269 TAAAGAGAGCTCATTTAAATACCAATTAGTCTAAAAACAGGAGGTAAAGAAATA---AAT 1325

Qy 468 CATCTATTGCTGAGACGACAGCGGAGGACAGGATCGGATATATAA-CCCAAGGCATT 526
Db 1326 CATCTCTCGCTGACAGCAGAGGAAAGGGGCAATGATTGGGATATATAAATCCCAAGGCATT 1385

Qy 527 CGAGCCGGCA-ACGGCAACCCCTTTGGGTCCCTCCCTCTTTGTATGGCGCTCTGTTTTC 585
Db 1386 CGAGCTGGAGTGGGCAACCCCTTTGGGTCCCTCCCTCTTTGTATGGAGCTCTGTTTTC 1445

Qy 586 ACTCTATTCTACTCTATTAATCTTGCACTTGAACTGAAAAAATATAAATAAATAA 635
Db 1446 ACTCTATTCAATCTTGCACTTGCACTGATATAAATAAATAAATAAATAAATAA 1495

RESULT 19
AG033781/c 653 bp DNA linear GSS 01-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-008021.F, genomic survey sequence.
DEFINITION AG033781
ACCESSION AG033781
VERSION AG033781.1 GI:16560654
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
JOURNAL Totoki, Y., Watanabe, H. and Sakaki, Y.
REFERENCE BAC end sequences of Library PTB
AUTHORS Unpublished
TITLE 2 (bases 1 to 653)
JOURNAL Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
AUTHORS Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
    Location/Qualifiers
        1..653
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-008021.F"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

```

Query Match 66.6%; Score 423; DB 9; Length 653;
 Best Local Similarity 88.0%; Pred. No. 3e-111;
 Matches 51; Conservative 0; Mismatches 52; Indels 18; Gaps 4;
 2 CCTGTATCTTAACTCC-TTGTTAAGTTGTCTCTCCAGAAATCAAACTGTAAACTA 60
 639 CTTGTTCTTAAACCCCGTGTAAAGTGTCTCTCCAGAAATCAAGCTGTAAAGCTA 580
 61 CAAATGTTCTTCAAAATGGAGCAGCAGTGGAGTCCATGACTAAGATCCACCGTGGACCC 120
 579 CAAATGTTCTTCAAAATGGAGCAGCAGTGGAGTCCATGACTAAGATCCACCGTGGACCC 535
 121 CTGACCGGCTCTAGCCCATGCTCGATGTTTAATGACATTTGAAGGCACCCCTCCCGAG 180
 534 CTGACCGGCTCTAGCCCATGCTCGATGTTTAATGACATTTGAAGGCACCCCTCCCGAG 476
 181 GAAATCTCAACTGACACACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGAGGGT 240
 475 AAAATCTCAACTGACACACCCCTACTATGCCCCAAATTCAGCGGGAAGCAGTTAGAGGGT 416
 241 CATCAGCAACCTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 300
 415 CGTTGGCCAACTCCCAACAGCAGTGGGTTTCTGTTGAGAGGGGAGCTGAGAGAC 356
 301 AGGACTAGCTGGATTCTCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 355 AGGACTAGCTGGATTCTCTAGGCCAACGAAAGATCCCTAAGCCTAGCTGGGAAGGTGACT 296
 361 GCATCCACCTCTAAACATGGGGTTGCAACTTAGCTCACACCCGACCAATCA-GAGAGCT 419
 295 ACATCCACCTTTAAACATGGGGTTGCAACTTAGCTCACACCCGACCAATCA-GAGAGCT 236
 420 CACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCCT 479
 235 CACTAAATCTTAATAGGCAAAATAGGAGGTAAAGAAATAGCCATCATCTATTGCCT 176
 480 GAGAGCACAGCGGAGGAGCAAGATCGGATATAAACCCAGGATTCGAGCGGCAACG 539
 175 GAGAGCACAGCGGAGGAGCAAGATCGGATATAAACCCAGGATTCGAGCGGCAACG 116
 540 GCAACCCCTTTTGGGTCCTCCCTCTTGTATGGCGCTCTG 580
 115 GCTACCTCTTGGGTCCTCCCTCTTGTATGGAGCTCG 75

RESULT 20
 AI598135/c
 LOCUS
 DEFINITION
 tnl410.xl NCI CGAP Brn25 Homo sapiens cDNA clone IMAGE:2167578 3'
 similar to contains_PFR5.tl PFR5 repetitive element ;, mRNA
 sequence.
 AI598135
 AI598135.1 GI:4607183
 EST.
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 490)
 NCBI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute / National Institute of Neurological
 Disorders and Stroke, Brain Tumor Genome Anatomy Project
 (CGAP/BTGAP), Tumor Gene Index
 Unpublished (1998)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
 Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1407 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 455
 POLYA=No.

FEATURES
 source
 1..490
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2167578"
 /tissue_type="anaplastic oligodendroglioma"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Brn25"
 /note="Organ: brain; Vector: pTV73D-Pac (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I oligo(dT) primer [5'
 TGTTACCAATCTGAATGGAGCGGCGCATAGGTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pTV73 vector.
 Library is normalized, and was constructed by Bento
 Soares and M.Fatima Bonaldo."

ORIGIN
 Query Match 65.5%; Score 415.8; DB 1; Length 490;
 Best Local Similarity 92.9%; Pred. No. 3.5e-109;
 Matches 435; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
 126 CCGCGCTCTAGCCCAATGCTCCGATCTTAATGATGAAGGACCCCTCCCGAGAAAT 185
 490 CCAGCCTCTAGCCCTTGTTCGATGTTAATGATCAAAAGGACCCCTCTCGAGAAAT 431
 186 CTCACCTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGTCA 245
 430 CTCANCTGCACACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGTCA 371
 246 GCCAACCTCCCAACAGCAGCTTGGGTTTCTCTGAGAGGGGGGACTGAGAGCAGGAC 305
 370 GTCAACCTCCCAACAGCAGCTTGGGTTTCTCTGAGAGGGGGGACTGAGAGCAGGAC 311
 306 TAGCTGGATTCTCTAGGCCAAACGAGAATCCCTAAGCCTAGCTGGGAAGTGTGATC 365
 310 TAGCTGGATTCTCTAGGCCGACTAAGAATCCCTAAGCCTAGCTGGGAAGTGTGATC 251
 366 CACCTCTAAACATGGGGCTTGCACCTAGCTCACACCCGACCAATCAGAGAGCTCACTAA 425
 250 CACCTTTAAACACGGGGCTTGCACCTAGCTCACACCCGACCAATCAGAGAGCTCACTAA 191
 426 AATGCTAAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGC 485
 190 AATGCTAAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCCTGAGAGC 131
 486 ACAGCGGGAGGACAAAGGATCGGGATATAAACCCAGCAGCTTCGAGCGGCAACGCAACC 545
 130 ACAGCAAGAGGAGCAATGATCGGGATATAAACCCAGCAGCTTCGAGCGGCAACGCTACC 71
 546 CCCTTTGGGTCCTCCCTTTGTATGGCGCTCTGTTTTCACCTCTATT 593
 70 CTCTTTGGGTCCTCCCTTTGTATGGAGCGCTGTTTTCACCTCTATT 23

RESULT 21
 BE732673
 LOCUS
 DEFINITION
 601571305F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3925728 5',
 mRNA sequence.
 BE732673
 BE732673.1 GI:10146665
 EST.
 Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 440)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM752 row: j column: 01
High quality sequence stop: 440.
Location/Qualifiers
1..440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3925728"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

FEATURES
source
Query Match 62.6%; Score 397.4; DB 2; Length 440;
Best Local Similarity 94.1%; Pred. No. 7 4e-104;
Matches 413; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 179 AGGAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCG 238
DB 1 AGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCGGGAAGCAGTTAGAGCG 60
QY 239 GTCATAGCCCACTCCCAACAGCACTTGGGTTTCTCTGTGAGAGGGGCACTGAGAG 298
DB 61 GTCGTCGGCACTCCCAACAGCACTTGGGTTTCTCTGTGAGAGGGGCACTGAGAG 120
QY 299 ACAGGACTAGCTGGATTCTTAGGCCAAACGAAGATCCCTTAAGCCTAGCTGGGAAGGTGA 358
DB 121 ACAGGACTAGCTGGATTCTTAGGGCTGACTAAGAAATCCCTAAGCCTAGCTGGGAAGGTGA 180
QY 359 CTCGATCCACTCTAAACATPGGGGCTTGCACCTTAGCTACACCCGACCAATCAGAGAGC 418
DB 181 CCACATCCACTTAAACAGCGGCTTGCACCTTAGCTACACCTGACCAATCAGAGAGC 240
QY 419 TCATAAAATGCTAATTAGCCAAATAGGAGCTAAGAAATAGCCAATCATCTATTGCC 478
DB 241 TCATAAAATGCTAATTAGCCAAAGACAGAGGTAAGAAATAGCCAATCATCTATTGCC 300
QY 479 TGAGAGCAGCGGGAGGGAACAGGATCGGGATATAAACCCAGGCATTCGAGCCGCAAC 538
DB 301 TGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAACCCAGTCCTTCGAGCCGCAAC 360
QY 539 GGCAACCCCTTTGGGTCCTCCCTTTGATGGGGCTCTGTTTTCACCTCTATTTCATT 598
DB 361 GGCAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCACCTCTATTTCATT 420
QY 599 CTATTAAATCTTGCACACTG 617
DB 421 CTATTAAATCTTGCACACTG 439

RESULT 22
BI963185/c
LOCUS
DEFINITION i57c10.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5670930 3', mRNA sequence.
ACCESSION BI963185
VERSION BI963185.1 GI:16337590
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R.,
Williams, T., Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 451.
Location/Qualifiers
1..543
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5670930"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecot of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

FEATURES
source
Query Match 62.4%; Score 396; DB 4; Length 543;
Best Local Similarity 88.3%; Pred. No. 2e-103;
Matches 467; Conservative 0; Mismatches 51; Indels 11; Gaps 3;
QY 74 AAATGGAGCACAGATGGAGTCCATGACTAAGATCCACCGTGGACCCCTGGACCGGCTG 133
DB 543 AAATGGAGCCCCAGATGCGAGTCCATGACTAAGATCTACTGCGGAGCTCTGGACCGGCTG 484
QY 134 CTAGGCCATGCTCCGATGTTAATGACATTAAGGACACCCCTCCCGAGGAAATCTCAACTG 193
DB 483 CTAGGCCA-GCTCCACATTAATGACATCAAGGACACCCCTCCCGAGGAAATCTCAACTG 425

```
QY 194 CACAAACCCCTACTATGCCCAATTTACGCGGGAAGCAGTTAGAGGGGTTCATCAGCCAACT 253
Db 424 CATGACACCTACTACGCCCCCAATTCAGCAGAAAGCAGTTAGAGTGGTCTGGCCAACT 365
QY 254 CCCCAACAGACACTTGGGTTTTCTCTGTGAGAGGGGGGACTGAGAGACAGACTAGCTGGA 313
Db 364 CCCCAACAGACACTTGGGTTTT-CTGTTGAGATGGGGGACTGAGAGACAGACTAGCTGGA 306
QY 314 TTTCTTAGGCAAGAAATCCCTAGCTAGCTAGCTGGGAAGGTGACTGCTACACCTCTA 373
Db 305 TTTCTTAGGCAACTAAGAAATCCCTAAGCTAGCTGGGAAGGTGACTTACACCCACCTTTA 246
QY 374 AACATGGGCTTGCAACTTAGCTACACACCCGACCAAT-----CAGAGAGCTCACTA 424
Db 245 ACNACTAGGCTTGCAACTTAGCTACACACCCGACCAANTAGGAGTAGAAGAGAGCTTGCTA 186
QY 425 AAATGCTAATTAGGCAAAATAGAGGTAAAGAAATAGCCAAATCATCTATTGCTGAGAG 484
Db 185 AAATGCTAATTAGGCAAAACAGAGGTAAAGAAATAGCCAGTCATCTATCGCTGACAG 126
QY 485 CACAGCGGAGGACAGAGTTCGGGATATAAACCAGGCAATTCAGCGGCAACGCAAC 544
Db 125 CACAAGGGGGGGCAATATGATCAGGATATAAATCAAGCTCAGGCAATTCAGGCAATGGCTAC 66
QY 545 CCCCTTTGGGTCCCTCCCTCTGTTGATGGGCTCTGTTTTCACCTCTATT 593
Db 65 CCACCTTTGGGTCCCTCCCTCTGTTTATGGGAGCTCTGTTTTCACCTCTATT 17
```

RESULT 23

BX380176/c
LOCUS BX380176 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI044YK06 3-PRIME, mRNA sequence.

ACCESSION

VERSION BX380176.2 GI:46834414

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 966)

Li W.B., Gruber C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 8, 2003 this sequence version replaced gi:30460243.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CSODI044BF03NP1&c=4215.r.

FEATURES

source

1..966
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI044YK06"
/tissue="PLACENTA"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```
Query Match 62.2%; Score 395.2; DB 5; Length 966;  
Best Local Similarity 79.6%; Pred. NO. 3.8e-103;  
Matches 465; Conservative 22; Mismatches 84; Indels 13; Gaps 2;  
QY 1 CCTGTATCTTTAAACCTCCTCTGTTAAAGTTTGTCTTCCAGAAATCAAAACTGTAAACTA 60  
Db 574 CCTGTATCTTTAAACCTCCTCTGTTAACTTTGTCTTCCAGAAATCAAGCTGTAAACTA 515  
QY 61 CAAATTTGTTCTTTCAATGGAGCACAGATGGAGTCCATGACTTAAGATCCACCGTGAGACC 120  
Db 514 -----CAAATGGAGCCARGATGCAGTCAGTCAGCTACCGCGGCC 467  
QY 121 CTGAGACCGGCTGTAGCCCATGCTCCGATGTTTAATGACATTTGAAGGCCACCTCTCCGAG 180  
Db 466 CTGAGACCGGCTGTAGCCCGCCGCTGATGTTTAATGACATCAAAAGGCCACCTCTCCGAG 407  
QY 181 GAAATCTCAACTGACCAACCCCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGT 240  
Db 406 GCAATCTCAGCTGCAACCCCTCTACTACGCCCAATTCGCGGAAGCAGTTAGAGCGGT 347  
QY 241 CATCAGCCAACTCCCAACAGCAGCTTTGGGTTTCTGTTGAGAGGGGGGACTTGAGAGAC 300  
Db 346 CGTCGCCCAACCTCCCGACGACCTTGGGTTTCTGTTGAGATGGGGGACTTGAGAGAC 287  
QY 301 AGGACTAGCTGGAATTTCTAGGCCAAAGAAATCCCTTAAGCCTAGCTGGGAAGGTGACT 360  
Db 286 MGGACTGGCTGGATTTCTTAGGCTGACTAAGAAATCCCTTARGCCTRGCTGGGAAGGTGACC 227  
QY 361 GCATCCACCTCTAAACATGGGGCTTGCACCTTAGCTCACACCCGACCAATCAGAGAGCTC 420  
Db 226 ACATCCACCTTTAAACMCGGGGCTTGCCACTTGGCTCACCCCTGCCCAATTCRGGGGCTC 167  
QY 421 ACTAAATGCTAATTAGGCAAAATAGGAGGTAAAGAAATAGC-CAATCATCTATTGCTCT 479  
Db 166 ACTAAATGCTAATYCGGCCCCSCSSSSSVSSSSGCCCTCGCGMSYCMTCATTGCTCT 107  
QY 480 GAGAGCACGCGGGAGGACAGGATCGGATATAAACCCAGGCATTCGAGCGGCAACG 539  
Db 106 GAGAGCACAGCMGAGGAGCAATGATCGGATATAAACCCAMGTCTTCGCGCGGCCCG 47  
QY 540 GCACACCCCTTTGGGTCCCTCCCTTGTATGGCGCTCTGTTT 583  
Db 46 GCCCCCCCTTTGGGTCCCTCCCTTTCGTTNNTGGGAGCTCTGT 3
```

RESULT 24

AG058970/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AG058970 651 bp DNA linear GSS 02-NOV-2001
Pan troglodytes DNA, clone: PTB-046A08.R, genomic survey sequence.

AG058970 1 GI:16596431

GSS.

Pan troglodytes (chimpanzee)

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library PTB

2 (bases 1 to 651)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail: chimpbes@gsic.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/.

Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end

was generated during the R&D process and may have higher chance of

clone tracking errors.


```

PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .651
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-04A08.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
ORIGIN
Query Match 62.2%; Score 394.8; DB 9; Length 651;
Best Local Similarity 87.6%; Pred. No. 4.6e-103; Indels 14; Gaps 6;
Matches 503; Conservative 0; Mismatches 57;

QY 20 TTGTTAAGTTTGTCTTCCCA-GAATCAAAACTGTAAACTACAAATTTGTTCTCAAAATG 78
DB 613 TTGATAGTTTGTCTCTCCAGATCAAGCTGTAAACTACAAATGGTTCTTCAAAATG 554

QY 79 GAGCACAGATGAGTCCATGAC-TAAGATCCACGCTGAGACCCCTGGACCGCCTGCTAG 137
DB 553 GAGCACAGATACAGTGCATGTCTTAAGTTCTCCCGTGAACCCCTGGATGGCCAGTAG 494

QY 138 CCATGCTCGATGTTAATGACATTCGAAGCCACCCCTCCGAGGAAATCTCAACTGCACA 197
DB 493 CCATGCTCGATGTTAATGACATTCGAAGCCACCCCTCCGAGGAAATCTCAACTGCACA 434

QY 198 ACCCTCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGTCTATCAGCCAACTCCCC 257
DB 433 ACCCTCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGTCTATCAGCCAACTCCCC 374

QY 258 AACAGCACTTGGGTTTCTTGTGAGAGGGGACGACTGAGACAGGACTAGCTGGATTTC 317
DB 373 AACAGCACTTGGGTTTCTTGTGAGAGGGGACGACTGAGACAGGACTAGCTGGATTTC 314

QY 318 CTAGGCGCAAGAAATCCCTAAGCCTAGCT-GGGAAGGTGACTGCATCCACTTAAAC 376
DB 313 CTAGGCTGACTAAGATTCCTAAGCCTAGCTGGGGAAGGTGATTGCACCACTTAAAC 254

QY 377 ATGGGGCTTGCAACTTAGCTCACACCGACCAATCAG-----AGAGCTCACTAAA 427
DB 253 ATGGGGCTTGCAACTTAGCTCACACCGACCAATCAGTTAGTAAAGAGGGCTCACTAAA 194

QY 428 TGCTAATTAGCC-AAAATAGGAGGTAAAGAAATAG-CCAATCATCTATTGCTTGAGAGC 485
DB 193 TACAAATTAGCTTAAAGAGGAGATTAAGAAATAGTCAANTCATATACCTGAGACC 134

QY 486 ACAGCGGGAGGACAAGGATCGGATATAAACCCAGGCATTCGAGCGGCAACGGCAACC 545
DB 133 ACAGTGGGAGGACAATGGTTCGGGTATAAACCCAGGCATTCGAGCAGGAGTGGCAACC 74

QY 546 CCCTTTGGGTCCCTCCCTTTTGTATGGCGCTCT 579
DB 73 CGCTTTGGGTCCCTCCCTTTGTTATGGGAGCTCT 40

RESULT 25
AG126669/c
LOCUS
DEFINITION Pan troglodytes DNA, clone: PTB-137E19.F, genomic survey sequence.
ACCESSION AG126669
VERSION AG126669.1 GI:16655834
KEYWORDS GSS.
SOURCE GSS.
ORGANISM Pan troglodytes (chimpanzee)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 701)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chuimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .701
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-137E19.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 61.9%; Score 393.2; DB 9; Length 701;
Best Local Similarity 88.1%; Pred. No. 1.4e-102; Indels 10; Gaps 3;
Matches 465; Conservative 0; Mismatches 53;

QY 79 GAGCACAGATGGAGTCCATGACTTAAGATCCACGCTGAGACCCCTGGACCGCCTGCTAGC 138
DB 651 GACCCGAGATCCAGTCCATGACTTAAAGACTACTGTGAACCCCTGGATGGCCTGCTAGC 592

QY 139 CCATGCTCCGATGTTAATGACATTTGAAGGACCCCT-CCGAGGAAATCTCAACTGCACA 197
DB 591 CCATGCTCCGATGTTAATGACATTTGAAGGACCCCTGCCCCGAGGAAATCTCAACTGCACA 532

QY 198 ACCCTCTACTATGCCCAATTCAGCGGGAAGCAGTTAGAGCGGTCTATCAGCCAACTCCCC 257
DB 531 ACCCTCTACTATGCCCAATTCAGCGGGAAGCAGTTAAGTGGTGGTGGCCAACTCCCC 472

QY 258 AACAGCACTTGGGTTTCTTGTGAGAGGGGAGCTAGAGACAGGACTAGCTGGATTTC 317
DB 471 AACAGCACTTGGGTTTCTTGTGAGAGGGGAGCTAGTGAAGAGGAGGACTAGCTGGATTTC 412

QY 318 CTAGGCCAACGAAAGTATCCCTAAGCTAGCTGGGAAGGTGACTGCATCCACTTAAACA 377
DB 411 CTAGGCTCAGCTAAGAAATCCCTAAGACTAGCTGGGAGGTGACCATCCACTTGAACA 352

QY 378 TGGGGCTTGGCAACTTAGCTCACACCCGACCAATC-----AGAGACTCACTAAAATG 429
DB 351 CGGGGCTTGGCAACTTAGCTCACATCCGACCAATCAGGTAAATAAGAAAGATAACTAAAATG 292

QY 430 CTAATTAGGCAAAAATAGAGGTAAGAAATAGCAATCATCTATTGCTTGAGAGCACAG 489
DB 291 CTAATTAGGCAAAAACAGAGGTAAGAAATAGCAATCATCTATCATCTGAGTACAG 232

QY 490 CGGAGGAGGACAGGATCGGATATAAACCCAGGCATTCGAGCGGCAACGGCAACCCCT 549
DB 231 CAGGAGGAGACATGATCGGGATATAAACCCAGGCATTCGAGCCAGCAACAGCTACCTCT 172

QY 550 TTGGGTCCCTCCCTCTTTGTATGGGCGCTCTGTTTTCACCTCTATTTCAC 597
DB 171 TT-GGTCCCTCCCTCTTTGTATGGGAGCTCTGTTTTCACCTCTATTAAAC 125
```

```

RESULT 26
AG066901/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H., and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 609)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H., and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpan@psc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..609
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-056L15.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 61.9%; Score 393; DB 9; Length 609;
Best Local Similarity 87.7%; Pred. No. 1.5e-102;
Matches 478; Conservative 0; Mismatches 55; Indels 12; Gaps 4;

Qy 47 AAACGTGAAACTACAAATTTCTTCAATGGAGCACCGATGGAGTCCATGACTAAGA 106
Db 606 AAGCTGTAGAACTACAAATCATTTCTTCAATGGAGCCCCAAATTCAGTCCACTAAGA 547
Qy 107 TCGACCGTGGACCCCTGGACCGCTCTAGCCCATGCTCC-GATGTTAATGACATTGAA 165
Db 546 TCTACTGCAGACCCCTTGACCGAGCTCTAGCTCATGCTCCAAATGTAATGGCCCTTGA 487
Qy 166 GGCACC-CCTCCGAGGAAATCTCAACTGCACACACCCCTACTATGCCCAATTCAGCGGG 224
Db 486 AGCCCTTCTCCGAGGAAATCTCACTGACACACCCCTACTACACCCCAATTCAGCAGG 427
Qy 225 AAGCAGTTAGACGGTCAATAGCAACCTCCCAACAGCACTTGGGTTTCTCTTTGAGA 284
Db 426 AAGCAGTTAGACGAGTCAATGCAACCTCCCAACAGCACTTGGGTTTCTCTTTGAGA 367
Qy 285 GGGGGACTGAGACAGGACTAGTGGATTCTCTAGCCCAACGAGATCCCTTAAGCCT 344
Db 366 TGGGGGACTTGAACAGGACTAGTGGATTCTCTA-GCTGACTAAGAAATCCCTAAGCCT 308
Qy 345 AGCTGGGAAGTGAATGATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACCCG 404

```

ORIGIN

Fri Feb 25 16:26:31 2005

FEATURES		Location/Qualifiers	
source		1..712	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/clones="Plate=3131 Col=8 Row=J"	
		/sex="male"	
		/clone_lib="CIT Approved Human Genomic Sperm Library D"	
		/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"	
ORIGIN			
Query Match		61.0%; Score 387.2; DB 8; Length 712;	
Best Local Similarity		82.2%; Pred. No. 7.5e-101;	
Matches 484; Conservative		0; Mismatches 94; Indels 11; Gaps 3;	
QY	184	ATCTCACTGCAACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTCA	243
DB	443	ATCTCACTGCAACACCCCTACTATGCCCCCAATTCAGCGGGAAGCAGTTAGAGCGGTCA	384
QY	244	CAGCCAACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGAGTCAAGACAGG	303
DB	383	CAGCCAACTCCCAACAGCAGCTTGGGTTTCTGTTGAGAGGGGGAGTCAAGACAGG	324
QY	304	ACTAGCTGGATTTCCTAGGCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGTGA	363
DB	323	ACTAGCTGGATTTCCTAGGCCAAATTAAGAATCCCTAAGCCTAGCTGGGAAGTGA	264
QY	364	TCCACTCTTAACATGGGCTTGCACCTAGCTTGCACCCGACCAATCAGAGAGCTCACT	423
DB	263	TCCACTCTTAACATGGGCTTGCACCTAGCTTGCACCCGACCAATCAGAGAGCTCACT	204
QY	424	AAATCTCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCCTGAGA	483
DB	203	AAATCTCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCCCTGAGA	144
QY	484	GCACAGCGGAGGGAACAGATCGGATATAACCCAGGATTCAGCGCGGCAACGGCAA	543
DB	143	GCACAGTGGGAGGGAACAGATTCGAATATAACCCAGGATTCAGCGCGGCAACGGCAA	85
QY	544	CCCCCTTTGGGTCCTCCCTTTGATGGCGCTCTGTTTTCACCTATTTCACCTATT	603
DB	84	CCGCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTTCACCTATTTCACCTATT	25
QY	604	AAATCTTGCACTGAAAAAATAA 627	
DB	24	AAATCTTGCACTGAAAAAATAA 1	
RESULT 28			
AQ892947/c		712 bp DNA linear GSS 10-NOV-1999	
LOCUS		HS 3131 B2 E04 MR CIT Approved Human Genomic Sperm Library D Homo	
DEFINITION		sapiens genomic clone Plate=3131 Col=8 Row=J, genomic survey	
ACCESSION		AQ892947	
VERSION		AQ892947.1	
KEYWORDS		GSS.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		1 (bases 1 to 712)	
		Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,	
		Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and	
		Hood,L.	
TITLE		Sequence-tagged connectors: A sequence approach to mapping and	
		scanning the human genome	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)	
MEDLINE		99380589	
COMMENT		1049764	
		Contact: Mahairas GG, Wallace JC, Hood L	
		High Throughput Sequencing Center	
		University of Washington	
		401 Queen Anne Avenue North, Seattle, WA 98109, USA	
		Tel: (206) 616-3618	
		Fax: (206) 616-3887	
		Email: jwallace@u.washington.edu	
		Clones may be purchased from Research Genetics (info@resgen.com).	
		BAC end Web Server: http://www.htsc.washington.edu	
		Plate: 3131 row: J column: 8	
		Seq primer: M13 Reverse	
		Class: BAC ends	
		High quality sequence stop: 712.	
RESULT 29			
AQ128526/c		436 bp mRNA linear EST 27-OCT-1998	
LOCUS		qc61h10.x1 Soares_placenta_8to9weeks_2NBHP8C09W Homo sapiens cDNA	
DEFINITION		clone IMAGE:1714147 3', similar to contains PTR7.bl PTR5 repetitive	
		element ;, mRNA sequence.	
ACCESSION		AQ128526	
VERSION		AQ128526.1	
KEYWORDS		GI:3597040	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		1 (bases 1 to 436)	
		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	
		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	
		EST.	
		Homo sapiens	
		qc61h10.x1 Soares_placenta_8to9weeks_2NBHP8C09W Homo sapiens cDNA	
		clone IMAGE:1714147 3', similar to contains PTR7.bl PTR5 repetitive	
		element ;, mRNA sequence.	
		GI:3597040	
		Homo sapiens (human)	
		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		1 (bases 1 to 436)	
		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	
		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	

AI393478.1	GI:4223025
EST.	
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 446)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 1061 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 445.
FEATURES	Location/Qualifiers
source	1..446 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:2111766" /lab_host="DH10B" /clone_lib="Soares_NFL_T_GBC_S1" /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBH19W, testis NHT, and B-cell NCI CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
ORIGIN	
Query Match	60.7%; Score 385.2; DB 1; Length 446;
Best Local Similarity	92.5%; Pred. No. 2.6e-100;
Matches	419; Conservative 0; Mismatches 24; Indels 10; Gaps
QY	183 AATCTCAACTGCACAAACCCCTACTATGCCCCCAATTTCAGCGGAGCAGTATTAGCGGTCA 241
DB	446 AATTTCAACTGTGCAACCCCTANTATGCCCAATTTCAGCAGGAGCAGTATTACGCGGTCA 387
QY	243 TCAGCCAACTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGACAG 300
DB	386 TCAGCCAACTCCCAACAGCAGCTTGGGTTTTCTGTTGAGAGGGGGAGCTGAGAGACAA 321
QY	303 GACTAGCTGGATTTCTTAGGCCCAACGAGAAATCCCTAAGCTAGCTGGAGGTGACTGC 361
DB	326 GACTAGCTGGATTTCTTAGGCCCGACTAAGAAATCCCTAAGCTAGCTGGAGGTGACTGC 266
QY	363 ATCCACCTCTAAACATGCGGCTTGCACCTTAGCTCACACCCGCAATTCAGAGAGCTCAC 421
DB	266 ATCCACCTTTAAACATGCGGCTTGCCACTTAGCTCACACCCCAATTCAGAGAGCTCAC 201
QY	423 TAAATGCTTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCAATCATCTATTGCTGAG 481
DB	206 TAAATGCTTAATTAGGCAAAAACAGGAGGTAAAGAAATAGCCAATCACCTTTTGCCTGAG 14
QY	483 AGCAGCGGGAGGACAGGATCGGATATAAACCAGGATTCGAGCGGCAACGCA 54
DB	146 AGCAGCAGGAGGAGCAGGATCGGATATAAACC CGGGCATTTCAAGCCGGCAATGGCA 87
QY	543 ACCCCCTTTGGGTCCCTCCCTTTGATGGCGCTCTGTTTCTACTCTATTTCATCTCTAT 60
DB	86 ACCCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGT-----TTTCACTCTAT 37
QY	603 TAAATCTTGCAACTGAAAAA 635

Tumor Gene Index	
Unpublished (1997)	
Contact: Robert Strausberg, Ph.D.	
Email: cgapbs-r@mail.nih.gov	
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.	
Insert Length: 720 Std Error: 0.00	
Seq primer: -40m13 fwd. RT from Amersham	
High quality sequence stop: 428.	
FEATURES	Location/Qualifiers
source	1..436 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:171417" /dev_stage="two placentae: one from 8 weeks and another from 9 weeks post conception" /lab_host="DH10B (ampicillin resistant)" /clone_lib="Soares placenta 809weeks 2NBHP8to9W" /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN	
Query Match	60.9%; Score 386.8; DB 1; Length 436;
Best Local Similarity	93.7%; Pred. No. 8.8e-101;
Matches	403; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY	206 TATCCCCCAATTTCAGCGGAGCAGTATTAGAGCGGTCTATCAGCCAACTCCCAACAGCAC 265
DB	436 TACGCCCAATTTCAGCAGGAGCAGTTAGAGCGGTCTCGGCCAACTCCCAACAGCAC 377
QY	266 TTGGGTTTTCTGTGAGAGGGGGAGTGTAGAGCAGAGCTAGCTAGTGTCTTAGGCCA 325
DB	376 TTAGGTTTTCTGTGAGTGGGGAGTGTGAAGACAGGAGCTAGCTAGTGTCTTAGGCTG 317
QY	326 ACGAAGATCCCTAAGCTAGCTGGGAGGTGAGTGCATCCACCTCTAAACATGGGGCTT 385
DB	316 ACTAAGAATCCCTAAGCTAGCTGGGAGGTGAGTGCATCCACCTTTAAACATGGGGCTT 257
QY	386 GCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAATGCTTAATTAGGCAAAAT 445
DB	256 GCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAATGCTTAATTAGGCAAGAC 197
QY	446 AGGAGGTAAGAAATAGCAATATCTATTGCTTGTAGAGCAGCGGGAGGAGCAAGAT 505
DB	196 AGGAGGTAAGAAATAGCAATATCTATTGCTTGTAGAGCAGCGGGAGGAGCAATGAT 137
QY	506 CGGATATTAACCCAGCATTCGAGCGGCAACCGCAACCCCTTTGGGTCCCTCCCTT 565
DB	136 CGGATATTAACCCAGCATTCGAGCGGCAACCGCAACCCCTTTGGGTCCCTCCCTT 77
QY	566 TGTATGGCGCTGTGTTTCACTCTATTTCATCTATTAAATCTTTCGCACTGAAAAAAA 625
DB	76 TGTATGGAGCTCTGTTTTCATGCTATTTCATTTTAAATCTTTGCAACTGCAAAAAA 17
QY	626 AAAAAAAA 635
DB	16 AAAAAAAA 7

RESULT 30	
AI393478/c	
LOCUS	446 bp mRNA linear EST 30-MAR-1999
DEFINITION	t945g04.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:	2111766 3', mRNA sequence.
AI393478	
ACCESSION	

[illegible]

Db 480 CACATGCTCGATGTTAATGACATCGAAGGACACCCGCTCTCGAGGAATCTCAACTGCACA 421
 QY 198 ACCCTACTATGCCCCCAATTCACGGGAAAGCAGTTAGAGGGTCATCAGCCACCTCCCC 257
 Db 420 ACCCTACTACACCCCAATTCACGGAAGCAGTTAGAGGGTCGTCAGCCACCTCCCC 361
 QY 258 AACAGCAGTCTGGGTTTCTGTTGAGAGGGGGGACTGAGAGCAGGACTAGCTGGATTTC 317
 Db 360 AGCAGCAGTCTGGGTTTCTGTTGAGAGGGGGGACTGAGAGCAGGACTAGCTGGATTTC 301
 QY 318 CTAGGCGCAAGAAATCCCTAAGCCTAGCTGAGAGGGGGGACTGAGAGCAGGACTAGCTGGATTTC 377
 Db 300 CTAGGCTGACTAAGAAATCCCTAAGCCTAGCTGAGAGGGGGGACTGAGAGCAGGACTAGCTGGATTTC 241
 QY 378 TGGGGCTTCAACTAGCTCACACCCGACCAATC-----AGAGAGCTCACTAAAT 428
 Db 240 CGGGGCTTCAACTAGCTCACACCCGACCAATCAGATAGTAAGAGAGCTCACTAAAT 181
 QY 429 GCTAATTAGGCAAAATAGGAGGTAAAGAAATAGCCCAATCATCTATTGCTGAGAGCACA 488
 Db 180 GCTAATTAGGCAAAACAGAGGTAGAGAAATAGCCCAATCATCTATCGCTGAGAGCACA 121
 QY 489 GCGGGAGGCAAGGATCGGATATAAACCCAGGCTTCGAGCGGCAACGGCAACCCCC 548
 Db 120 GCAGGAGGCAATGATCCGATATAAACCAAGCATTCGAGCGGCAATGGCTACCCCTC 61
 QY 549 TTTGGTCCCTCCCTTTGATGGCGCTCTGTTTCACTTATT 593
 Db 60 TTTGTGTCCTCCCTTTGATGGAGCTCTATTTTCACTTATT 16

RESULT 33
 BX481837/c
 LOCUS
 DEFINITION
 DXFZp686M05227 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 DXFZp686M05227 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 664)
 Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
 Fobo, G., Han, M. and Wiemann, S.
 EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
 Unpublished (2003)
 Contact: MIPS
 MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No sl sequence available.
 This clone (DKFZp686M05227) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

FEATURES
 source
 1..664
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686M05227"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hlcc3)"
 /note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
 cDNA-collection"

ORIGIN

Query Match 59.8%; Score 379.8; DB 5; Length 664;
 Best Local Similarity 84.3%; Pred. No. 1e-98;
 Matches 514; Conservative 0; Mismatches 78; Indels 18; Gaps 7;
 QY 40 AGAATCAAACTGTAACAACTACAAATGTTCTTTCAATGGAGCAGAGTGGAGTCCATG 99
 Db 664 AGAATCAAACTGTAACAACTACAAATGTTCTTTCAATGGAGCAGAGTGGAGTCCATG 605
 QY 100 ACTAAGATCCACCTGGACCCCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGAC 159
 Db 604 ACTAAGATCCACCTGGACCCCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGAC 545
 QY 160 ATTGAAGGACCCCTCCCGAGGAATCTCACTGACAAACCCCTACTATGCCCAATTC 219
 Db 544 ATCGAAGGACCTCTCCCGAGGAATCTCAATGCAACCCCTACTATGCCCAATTC 485
 QY 220 GCGGGAAGCAGTGTAGAGCGGTCTCAGCCAACTCCCAACAGCAGCTTCGGTTCCTGT 279
 Db 484 GCGGGAAGCAGTGTAGAGCGGTCTCAGCCAACTCCCAACAGCAGCTTCGGTTCCTGT 425
 QY 280 TGAGAGGGGGGACTGAGAGCAGGACTAGCTGGATTCTTAGGCCCAACGAAGATCCCTA 339
 Db 424 TGAGAGGGGGGACTGAGAGCAGGACTAGCTGGATTCTTAGGCCGACTAACAATTTCTA 365
 QY 340 AGCTTAGCT-GGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
 Db 364 AGCTTAGTGGGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
 QY 397 CACACCCGACCAATC-----AGAGAGCTCACTATAATGCTAATAGGC-AAAAATA 446
 Db 304 CACATCTGCCCAATCAGGTAGTAAGAGAGAGCTATTAAATATCAATATTAGTCTTAAACA 245
 QY 447 GGAGGTAAAGAAATAGCCCAATCATCTATTGCTGCTGAGAGCAGCAGCGGAGGAGCAAGATC 506
 Db 244 GGAGGTAAAGAAATA---AATCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 188
 QY 507 GGGATATAAA-CCAGGCAATTCAGCGGCA-ACGGCAACCCCTTTGGGTGCTCCCTCCCT 564
 Db 187 GGGATATAAAACCCAGGCAATTCAGCTGGAGTGGGCAACCCCTTTGGGTGCTCCCTCCCA 128
 QY 565 TTGTATGGGGCTCTGTTTCACTCTATTTCATCTATTAAATCTTGCACACTGAAAAA 624
 Db 127 TTGTATGGGAGCTCTGTTTCACTCTATTTCATCTATTTCATCTGCAACTGAAAAA 68
 QY 625 AAAAAA 634
 Db 67 AAAAAA 58

RESULT 34
 AU158595/c
 LOCUS
 DEFINITION
 AU158595 PLACE3 Homo sapiens cDNA clone PLACE3000382 3', mRNA
 sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 556)
 Ota, T., Suzuki, Y., Saito, K., Ishii, S., Yamamoto, J., Sugiyama, T.,
 Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y. and Isogai, T.
 HRI human cDNA project (Ota, T., Suzuki, Y., Saito, K., Ishii, S.,
 Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
 Masuho, Y., Isogai, T.)
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975

Fri Feb 25 16:26:31 2005

Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1. .556
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE3300382"
/tissue_type="placenta"
/clone_lib="PLACE3"
/note="Vector: pME18SFL3"

FEATURES
source

ORIGIN
Query Match 59.1%; Score 375.2; DB 1; Length 556;
Best Local Similarity 84.1%; Pred. No. 2.1e-97; Indels 39; Gaps 3;
Matches 470; Conservative 0; Mismatches 50

QY 48 AACTGTAAATACAAATGTTCTTCAAAATGGAGCACCCAGATGGAGTCCATGACTAAGAT 107
DB 540 AACTGTAAATGTTCTTCAAAATGGAGCACCCAGATGGAGTCCATGACTAAGAT 482
QY 108 CCACCGTGGACCCCTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACATTGAAGG 167
DB 481 CTACNGTGNACCTTGGAGCTGGCTGTAGCCCATGTTCTGATGTTAATGACATCAAGG 422
QY 168 CACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGGAAG 227
DB 421 CATCCCTCCGAGGAATCTCAACTGCACAAACCCCTANTATGCCCAATTCAGCGAAG 362
QY 228 CAGTTAGAGCGGTATCAGCAACCTCCCAACAGACATTTGGTCTTCTGTTGAGAGG 287
DB 361 CAGTTAGAGCGGTATCAGCAACCTCCCAACAGACATTTGGTCTTCTGTTGAGAGG 302
QY 288 GGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCGGACAGAGATCCCTAAGCTAGC 347
DB 301 GGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCGGACAGAGATCCCTAAGCTAGC 242
QY 348 TGGGAGGTGACTGCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACC 407
DB 241 TGGGAGGTGACTGCATCCACCTCTAAACATGGGCTTGCAACTTAGCTCACACCCGACC 182
QY 408 AATC-----AGAGAGTCTACTAAATGCTTAATTAGGCAAAATAGAGGTAAGAA 458
DB 181 AATCAGGTAGCAGAGAGCTGTTAAATGATAATTAGGCAAAATAGAGGTAAGAA 122
QY 459 ATAGCCCAATCTATTGCTGAGACACAGCGGAGGACAGGATCGGGATATAAACC 518
DB 121 ATAGCCCAATCTATTGCTGAGACACAGCGGAGGACAGGATCGGGATATAAACC 91
QY 519 CAGGCAATTCAGCGCGCAACGGCAACCCCTTTGGGTCCCTCTTGTATGGGCGCTC 578
DB 90 CAGGCAATTCAGCGCGCAACGGCAACCCCTTTGGGTCTCTCTTGTATGGGAGCTC 31
QY 579 TGTCTTCACTCTATTTCAC 597
DB 30 TGTCTTCACTCTATTTCAC 12

RESULT 35
CB231128/c
LOCUS
DEFINITION
IMAGE:6884154 5', mRNA sequence.
CB231128
ACCESSION
CB231128.1 GI:28282706
VERSION
EST.
KEYWORDS
Macaca mulatta (rhesus monkey)
SOURCE
Macaca mulatta
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
1 (bases 1 to 865)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-r@mail.nih.gov
Tissue Procurement: Dr. Eliot Spindel
cDNA Library Preparation: CLONTECH
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM3133 row: i column: 17
High quality sequence stop: 373.
Location/Qualifiers
1. .865
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
/clone="IMAGE:6884154"
/tissue_type="ovary"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD Rh Ovi"
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.0-4.0 kb. Tissue pooled from
pre-pubertal, post pubertal sn menopausal monkeys.
Constructed by Clontech. Note: this is a NICHD Library."

ORIGIN
Query Match 58.2%; Score 369.6; DB 6; Length 865;
Best Local Similarity 87.1%; Pred. No. 9.9e-96;
Matches 458; Conservative 0; Mismatches 49; Indels 19; Gaps 4;

QY 46 AAAAAGTGTAAATACAAATGTTCTTCAAAATGGAGCACCCAGATGGAGTCCATGACTAAG 105
DB 516 AAGCTGTAAAGCTTACAAAGGTTCTCGAAGAGGAGGCGCCGATGCGAGTCCAGACTAAA 457
QY 106 ATCCACCTGTGACCCCTGGACCGGCTGCTAGCCATGCTCGATGTTAATGACATTGAA 165
DB 456 ATGTGCGGAGCA-CCTGTGAGTGGCTGCTAGCCATGCTCCG-----GACATCGAA 406
QY 166 GGCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGA 225
DB 405 GGCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGCGGA 346
QY 226 AGCAGTTAGAGCGGTATCAGCCAACTCCCAACAGACATTTGGGTTTTCTGTTGAGAG 285
DB 345 AGCAGTTAGATCAGTCATCAGTCACCTCCCAATAGACATTTGGGTTTTCTGTTGAG-G 287
QY 286 GGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCGCAAGAGAAATCCCTAAGCCTTA 345
DB 286 GGGTACTGAGAGACAGGACTAGCCGGATTTCTTAGGCGCAAGAGAAATCCCTAAGCCTTA 227
QY 346 GCTGGGAAGGTGACTGCTACCTCTAAACATGGGCTTGGCACTTAGCTCAGCCCGGA 405
DB 226 GCTGGGAAGGTGACTGCTACCTCTAAACATGGGCTTGGCACTTAGCTCAGCCCTGA 167
QY 406 CCAATC-----AGAGAGCTCTAAATGCTTAATTAGGCAAAATAGGAGGTTAAG 456
DB 166 CCAATCAGGTAGGAAAGAGAGATTCATATAATGCTTAATTAGGCAAAATAGGAGGTTAAG 107
QY 457 AATAGCCCAATCATCTATTGCTGAGAGCACAGCGGAGGAGCAAGGATCGGAGATATAAA 516
DB 106 AATAGCCCAATCATCTATTGCTGAGAGCACAGCGGAGGAGCAAGGATCGGAGATATAAA 47
QY 517 CCCAGGCAATTCAGCGCGGCAACCGCAACCCCTTTGGGTCCCTCC 562
DB 46 CCCAGGCAATTCAGCGCGGCAACCGCAACCCCTTTGGGTCCCTCC 1


```

RESULT 36
AI074704/c
LOCUS
DEFINITION
AI074704 470 bp mRNA linear EST 06-AUG-1998
ox83d05.s1 Soares_NhMpu.S1 Homo sapiens cDNA clone IMAGE:1662921
3' similar to contains PTR7.t2 PTR7 repetitive element ;, MRNA
sequence.
ACCESSION
AI074704.1 GI:3401348
VERSION
AI074704.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 470)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapsa-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 447.
FEATURES
Location/Qualifiers
1..470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1662921"
/tissue_type="pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NhMpu S1"
/notes="Forcan: mixed (see below); Vector: pRT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not 1;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NhMpu, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
ORIGIN
Query Match 57.6%; Score 365.8; DB 1; Length 470;
Best Local Similarity 90.8%; Pred. No. 1.1e-94;
Matches 415; Conservative 0; Mismatches 32; Indels 10; Gaps 2;

Qy 146 CCGATGTTAATGACATTAAGGACCCCTCCCGAGGAATCTCAACTGCACACCCCTAC 205
Db 470 CCGATGTTGATGACATCAAGGACCCCTCTCTGAGGAATCTCACTGCACGCCCTAC 411

Qy 206 TATGCCCAATTCAGCGGGAAGCATGTAGAGCGGTCTATGACCCAACTCCCAACAGCAC 265
Db 410 TATGCCCAATTCAGCGGGAAGCATGTAGAGCGGTCTATGACCCAACTCCCAACAGCAC 351

Qy 266 TTGGGTTTTCTGTGTGAGGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCCA 325
Db 350 TTGGGTTTTCTGTGTGAGGTGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTG 291

Qy 326 ACGAAGATCCCTAAGCCTAGCTGGGAAGGTGACTGTCATCCACCTCTAAACATGGGGCTT 385
Db 290 ACTAAGATCCCTAAGCCTAGCTGGGAAGGTGACGGCATCCACCTTTAAACACGGGCTT 231

Qy 386 GCAACTTAGCTCACACCCGACCAATC-----AGAGAGCTCACTAAATGCTAATTA 436
Db 230 GCAACTTAGCTCACACCCGACCAATCAGGGAGTAAAGACAGCTCACTAAATGCTAATTA 171

```

```

Qy 437 GCCTATATGAGGCTTAAGAAATAGCCAAATCATCTATTGCTTCCAGAGCAGACGGGAGG 496
Db 170 GGCTAAACAGGAGGTAAGAAATAGCCAAATCATCTATTGCTTCCAGAGCAGACGGGAGG 112

Qy 497 GACACGATCGGATATTAACCCAGGCAATTCAGCGCGCAACGGCAACCCCTTTGGGTC 556
Db 111 GACAATGATCAGATATTAACCCAGGCAATTCAGCGCGCAATTCAGCGCTACCGCTTGGGTC 52

Qy 557 CCTCTCTTTGATGGCGCTCTCTGTTTCACTCTATT 593
Db 51 CCTCTGTTTGTATGGGAGCTCTGTTTCACTCTATT 15

RESULT 37
AG076758/c
LOCUS
DEFINITION
AG076758 679 bp DNA linear GSS 03-NOV-2001
Pan troglodytes DNA, clone: PTB-071A23.R, genomic survey sequence.
ACCESSION
AG076758.1 GI:16628560
VERSION
AG076758.1
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
2 (bases 1 to 679)
Unpublished
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
LIBRARY
Location/Qualifiers
1..679
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-071A23.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 57.4%; Score 364.8; DB 9; Length 679;
Best Local Similarity 90.3%; Pred. No. 2.3e-94;
Matches 390; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 1 CCTGTATCTTTAACTCCCTTGTAGTTGTCTCTTCCAGAACTCAAACTGTAATACTA 60
Db 475 CCTGTATATTCACCTCTTGTAGTTGTCTCTTCCAGAACTTAAAGGCTGTAAGCTA 416

Qy 61 CAAATTTCTTCAAATGGAGCCAGATGGAGTCCATGACTAGATCCACCTGGGACCC 120
Db 415 CAAATGTTCTTCAAATGAAGCCAGATGCAGTCCATGGCTAAATCTACCATGTACCC 356

Qy 121 CTGACCGGGCTGCTAGCCCATGCTCCGATGTTTAATGACATTTGAGCGACCCCTCCGAG 180
Db 355 CTGGCGAGGCTGCTAGCCCTTGTCCGATGTTTAATGACATCAAAAGCAGCCCTCTCTGAG 296

```

QY 181 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAATTAGCGGGGAGCAGTTAGAGCGGT 240
 Db 295 GAAATCTCAACTGCACAAACCCCTACTATGCCCCAATTAGCGGGGAGCAGTTAGAGCGGT 236
 QY 241 CATCAGCNAACCTCCCAACAGCAGCTTGGGTTTTCTGTGAGAGGGGGAGCTAGAGAC 300
 Db 235 CGTCAGTCAACTCCCAACAGCAGCTTGGGTTTTCTGTGAGAGGGGGAGCTAGAGAC 176
 QY 301 AGGACTAGCTGGATTTCTTAGGCCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACT 360
 Db 175 AGGACTAGCTGGATTTCTTAGGCCCAACGAAGAATCCCTAAGCCTAGCTGGGAAGGTGACC 116
 QY 361 GATCCACCTCTAAACATGGGGCTTGCACCTTAGCTTACACCCGACCAATCAGAGAGCTC 420
 Db 115 GCGTCCACCTTTAAACACGGGGCTTGCACCTTAGCTTACACCCGACCAATCAGAGAGCTC 56
 QY 421 ACTAAATGCTA 432
 Db 55 TCGAGCATGCTA 44

RESULT 38
 A1128496/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 415)
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 This clone is available royalty-free through LML; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert length: 1184 Std Error: 0.00
 Seq primer: -40m13 fwd. Et from Amersham
 High quality sequence stop: 413.

FEATURES
 source
 1. .415
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1714118"
 /dev stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab host="DH10B (ampicillin resistant)"
 /clone lib="Soares_placenta_8to9weeks_2NbHP8to9W"
 /notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCGGCGGATTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M. Fatima Bonaldo."

ORIGIN
 Query Match 57.4%; Score 364.2; DB 1; Length 415;
 Best Local Similarity 93.7%; Pred. No. 31e-94;
 Matches 389; Conservative 0; Mismatches 25; Indels 1; Gaps 1;
 QY 203 TACTATGCCCAATTACGCGGGAAGCAGTTAGAGCGGTCTATCGCCCAACCTCCCAACAG 262

Db 415 TACTAGCCCCCAATTAGCAGGGAAGCAGTTAGAGCGGTCTCGGCCAA-CTCCNCAACAG 357
 QY 263 CACTTGGGTTTTCTGTGAGAGGGGGAGCTAGAGACAGGAGCTAGCTGGATTTCCTAGG 322
 Db 356 CACTTAGGTTTTCTGTGAGATGGGGGAGCTAGAGACAGGAGCTAGCTGGATTTCCTAGG 297
 QY 323 CCAACGAAGAATCCCTAAGCCTAGCTGGGGAAGGTGACTCCACTCTAAACATGGGG 382
 Db 296 CTGACTAAGAATCCCTAAGCCTAGCTGGGGAAGGTGACCATCCACCTTTAAACACGGG 237
 QY 383 CTTGCAACTTAGCTCACAACCCGACCAATCAGAGAGCTCCTAAATCTTAATAGGCAAA 442
 Db 236 CTTGCAACTTAGCTCACAACCTCAGCAATCAGAGAGCTCCTAAATCTTAATAGGCAAA 177
 QY 443 AATAGGAGGTAAAGAAATAGCAATCATCTATTTCCTGAGACACAGCGGAGGAGCAAG 502
 Db 176 AACAGGAGGTAAAGAAATAGCAATCATCTATTTCCTGAGACACAGCGGAGGAGCAAT 117
 QY 503 GATCGGATATAAACCCAGGCAATTCGAGCGGCAACCGGCAACCCCTTTGGGTCCCTCC 562
 Db 116 GATCGGATATAAACCCAGGCAATTCGAGCGGCAACCGGCAACCCCTTTGGGTCCCTCC 57
 QY 563 CTTTGTATGGCGCTCTGTTTTCACTCTATTTCCTCTATTAAATCTTGCAACTG 617
 Db 56 CTTTGTATGGAGCTCTGTTTTCACTCTATTTCCTCTATTAAATCTTGCAACTG 2

RESULT 39
 AG030228/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Pan troglodytes (chimpanzee)
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Tokoi, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB
 Unpublished
 2 (bases 1 to 777)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Tokoi, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimpbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS
 Sequencing: M13Rev
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.
 Location/Qualifiers
 1. .777
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-002120.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
 source

ORIGIN

Query Match	56.8%	Score	360.6	DB	9	Length	777		
Best Local Similarity	86.2%	Prod. No.	3.9e-93						
Matches	399	Conservative	0	Mismatches	64	Indels	0	Gaps	0

1	CCCTGTATCTTTAACTCCTTTGTTAAGTTTGCTCTTCCAGAAATCAAACTGTAAAACTA	60
515	CCCTATATCTTTAACTCCTTTGTTAAGTTTGCTCTTCTGGAATCAAAAGCTGCAAAACTG	456
61	CAAAATGTTCTTCAAAATGGAGCACAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC	120
455	CAAAATCAATCTTCAAAATGGAGCCCCAGATGCACTTATGACTAAGATCTTCCGTAGACCC	396
121	CTGGACCGGCTCTAGCCCCATGCTCCGATGTTAATGACATTTGAAGGCACCCCTCCCGAG	180
395	CTGGACCGGCTCTAGACCCATGCCCAATGTCAATGACATCGAAGGCACCCCTCCCGAG	336
181	GAATCTCAACTGCACAAACCCCTATCATATGCCCAATTCAGCGGGAGCAGTTAGAGCGGT	240
335	GAATCTCAATTCGACCAACCGGTAAATACGCCCAATTCAGCAAGAAGCAGTTTAGAGCAGT	276
241	CATCAGCCACCTTCCCCACAGCACATTTGGTTTCTGTGTAGAGGGGGAGCTCAGAGAC	300
275	CGTGGGCAACCTTCCCCAACAGCACATTTGGTTTCTGTGTAGAGCGGGGACTCAGAGAC	216
301	AGGACTAGCTGGATTTCTTAGGCCAACGAAGAATCCCTAAGCTAGCTGGGAAGGTGACT	360
215	AAGATTAGCTGGATTTCTTAGCCGACTAAGAAATCCCTAAGCTAGCTTGGAAAGGTGACT	156
361	GCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTTCACCCCGACCAATCAGAGAGCTC	420
155	GCATCCACCTTTAAACATGGGGCTTGTAACCTTAGCTTCCCAACCAACCAATCAAAAGAGCTC	96
421	ACTAAATGCTAATTAGGCAAAAATATAGGAGGTAAAGAAATGAC	463
95	TCGAGCATGCTAGCGGCGCCCTTATAGTAAGTCGTATATACAGC	53

RESULT 40
AI570707/c
LOCUS
DEFINITION
AI570707 linear EST 14-APR-1999
tm97g09.x1 NCI CGAP Brn25 Homo sapiens cDNA clone IMAGE:2164384 3',
similar to contains PTR7.t1 PTR7 repetitive element ;, mRNA
sequence.
AI570707
AI570707.1 GI:4534081
EST.
AI570707.1 Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 425)
NCI/NIHDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
cDNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1440 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 414.
Location/Qualifiers
1. .425
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned_image:2164384"
/cissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/cloned_lib="NCI CGAP_Brn25"
/mol_note="Organ: brain; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I- oligo(dT) primer [5', GTCTTACCACTGTAAGTGGAGCGCGCATAGTGTCTTTTTTTTTTTTTTTTCTTACCTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized, and was constructed by Bento Soares and M.Patima Bonaldo."

```

ORIGIN

Query Match	56.7%	Score 359.8	DB 1	Length 425
Best Local Similarity	93.3%	Pred. No. 5.9e-93		
Matches 376	Conservative 0	Mismatches 27	Indels 0	Gaps 0
191	CTGCACAACCCCTACTATGTCGCCCAATTACAGCGGGAGCAGTTAGAGCGGTCAATCAGCCAA	250		
425	CTGCACAACCCCTACCATGTCGCCCAATTACAGCAGGAAGCAGGTAGAGCGGTGTCAGTCAA	366		
251	CCTCCCAACAGCACTTGCGTTTTCTGTTGAGAGGGGGACTGAGAGACAGGACTAGCT	310		
365	CCTCCACAGCAGCACTTGCGTTTTCTGTTGAGAGGGGTACTGAGAGACAGGACTAGCT	306		
311	GGATTTCTTAGGCCCAACGAAGAATCCCTAAGCTAGCTGGGAAGGTGACTGCAATCACCT	370		
305	GGATTTCTTAGGCCGACTAAGAATCCCTAAGCTAGCTGGGAAGGTGACCGCGTCACCT	246		
371	CTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAATGC	430		
245	TTAAACACGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAATGC	186		
431	TAAATTAGCAAAAAATAGGAGTTAAGAAATAGCCAATCATCTATTGCGCTGAGAGCACAGC	490		
185	TAAATTAAAGCAAAAAACGAGAGTTAAGAAATAGCCAGTCACTATCGCGCTGAGAGCACAGC	126		
491	GGGAGGGACAGGATCGGGATATAAACCCACGGCATTCGAGCCGGCAACGGCAACCCCTTT	550		
125	AGAGGGACAAATGATCGGATATAAACCCAGGCATTCGAGCCAGCAACGGCTACCTCTTT	66		
551	TGGGTCCCTCCCTTTGATAGGGCGCTCTGTTTTCACTCTATT	593		
65	TGGGTCCCTCCCTTTGATGGAGCGCTGTTTTCACTCTATT	23		

Search completed: February 21, 2005. 14:51:02

Job time : 1986.67 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:32:54 ; Search time 6356.91 Seconds
(without alignments)
11288.850 Million cell updates/sec

Title: US-09-319-156B-9

Perfect score: 1481

Sequence: 1 agggccctccctatcatac.....gtcttacaatggaacccca 1481

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	1481	6	BD136197
2	1481	100.0	1481	6	AR344387
3	1481	100.0	1481	6	AX001027
4	1481	100.0	1932	14	AF127228
5	1468.2	99.1	162579	9	AL390039
6	1466.6	99.0	2055	6	CQ719381
7	1461.8	98.7	1629	6	BD252040
8	1461.8	98.7	1629	6	AX036992
9	1461.8	98.7	1629	14	AF331500
10	1461.8	98.7	172918	2	AC023366
11	1389.8	93.8	2074	6	AX478550
12	1341.8	90.6	10230	9	AY101589
13	1341.8	90.6	184675	9	AC145964
14	1340.2	90.5	10229	9	AY101586
15	1340.2	90.5	10229	9	AY101587
16	1340.2	90.5	10230	9	AY101588
17	1338.6	90.4	10246	9	AY101593
18	1338.6	90.4	10248	9	AY101592
19	1335.4	90.2	2694	9	AF520478

20	1335.4	90.2	2694	9	AF520487S2	AF520488 Homo sapi
21	1335.4	90.2	2694	9	AF520489S2	AF520490 Homo sapi
22	1335.4	90.2	2694	9	AF520505S2	AF520506 Homo sapi
23	1335.4	90.2	2694	9	AF520517S2	AF520518 Homo sapi
24	1335.4	90.2	2694	9	AF520519S2	AF520520 Homo sapi
25	1335.4	90.2	2694	9	AF520521S2	AF520522 Homo sapi
26	1335.4	90.2	2694	9	AF520529S2	AF520530 Homo sapi
27	1335.4	90.2	2694	9	AF520561S2	AF520562 Homo sapi
28	1335.4	90.2	2930	6	AX355872	AF525872 Sequence
29	1335.4	90.2	2930	9	AF208161	AF208161 Homo sapi
30	1335.4	90.2	2946	6	AR177269	AR177269 Sequence
31	1335.4	90.2	2946	6	BD232464	BD232464 Secreted
32	1335.4	90.2	2946	6	BD086133	BD086133 Secreted
33	1333.8	90.1	1624	9	AF506835	AF506835 Homo sapi
34	1333.8	90.1	1860	9	AF513360	AF513360 Homo sapi
35	1333.8	90.1	2055	6	BD221827	BD221827 Nucleic s
36	1333.8	90.1	2055	6	AX007999	AX007999 Sequence
37	1333.8	90.1	2599	6	BD221806	BD221806 Nucleic s
38	1333.8	90.1	2599	6	AX007978	AX007978 Sequence
39	1333.8	90.1	2694	9	AF520481S2	AF520482 Homo sapi
40	1333.8	90.1	2694	9	AF520483S2	AF520484 Homo sapi
41	1333.8	90.1	2694	9	AF520485S2	AF520486 Homo sapi
42	1333.8	90.1	2694	9	AF520491S2	AF520492 Homo sapi
43	1333.8	90.1	2694	9	AF520493S2	AF520494 Homo sapi
44	1333.8	90.1	2694	9	AF520495S2	AF520496 Homo sapi
45	1333.8	90.1	2694	9	AF520497S2	AF520498 Homo sapi

ALIGNMENTS

RESULT 1	BD136197	1481 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	BD136197	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.			
DEFINITION	BD136197	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses.			
ACCESSION	BD136197.1	GI:23231142			
VERSION	BD136197.1	GI:23231142			
KEYWORDS	JP 2002509437-A/7.				
SOURCE	unidentified				
ORGANISM	unclassified				
REFERENCE	1 (bases 1 to 1481)				
AUTHORS	Baccala,G.P., Pradel,F.K., Bedin,F., Sodoyer,M., Ott,C., Mallet,F., Perron,H. and Mandrand,B.				
TITLE	Retroviral nucleic material and nucleotide fragments, in particular, associated with multiple sclerosis and/or rheumatoid arthritis, for diagnostic, prophylactic and therapeutic uses				
JOURNAL	Patent: JP 2002509437-A 7 26-MAR-2002;				
COMMENT	BIO MERIEUX				
	OS Unidentified				
	PN JP 2002509437-A/7				
	PD 26-MAR-2002				
	PF 07-JUL-1998 JP 1999S08255				
	PR 07-JUL-1997 FR 97/08816				
	PI GLAUCIA PARAHNOS BACCALA,FLORENCE KOMURIAN PRADEL,FREDERIC PI BEDIN,				
	PI MIREILLE SODOYER,CATHERINE OTT,FRANCOIS MALLET,HERVE PERRON,				
	PI BERNARD MANDRAND				
	PC C12N15/48,C12Q1/70,C07K14/15,A61K31/70				
	CC Strandedness: Single;				
	CC Topology: Linear;				
	CC Retroviral nucleic material and nucleotide fragments, in CC particular,				
	CC associated with multiple sclerosis and/or rheumatoid arthritis,				
	CC for				
	CC diagnostic, prophylactic and therapeutic uses FH				
	CC Location/Qualifiers				
	FT source	1..1481			
	FT	/organism='Unidentified'.			
FEATURES	Location/Qualifiers				

source	1. .1481 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"			
ORIGIN	Query Match 100.0%; Score 1481; DB 6; Length 1481; Best Local Similarity 100.0%; Pred. No. 0; Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGCCCTCCCTATCATACACTTTTCTCTTTACTGTTCTCTTACCCCTTTTGGCTCTCACT	60	
DB	1	ATGGCCCTCCCTATCATACACTTTTCTCTTTACTGTTCTCTTACCCCTTTTGGCTCTCACT	60	
QY	61	GCACCCCTCCATGCTGTTACACACGATAGCTCCCTTACCAAGAGTTTCTATGAAGA	120	
DB	61	GCACCCCTCCATGCTGTTACACACGATAGCTCCCTTACCAAGAGTTTCTATGAAGA	120	
QY	121	ACGCGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC	180	
DB	121	ACGCGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC	180	
QY	181	ACCTTCACCTGCCACACCCATATGCCCGCACTGCTATACTCTGCCACTCTTTGCATG	240	
DB	181	ACCTTCACCTGCCACACCCATATGCCCGCACTGCTATACTCTGCCACTCTTTGCATG	240	
QY	241	CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT	300	
DB	241	CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT	300	
QY	301	GGAGCCACTGCTGTTGGACTTACTTACCCATACAGTATGCTCTGATGGGGTGAATT	360	
DB	301	GGAGCCACTGCTGTTGGACTTACTTACCCATACAGTATGCTCTGATGGGGTGAATT	360	
QY	361	CAAGTCAGCAAGAGAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT	420	
DB	361	CAAGTCAGCAAGAGAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT	420	
QY	421	AGCACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAACTACATGAACCCCTCCGTACC	480	
DB	421	AGCACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAACTACATGAACCCCTCCGTACC	480	
QY	481	CATAGTCCCTGCTGAGCCTATTATACACCTGCTCGGCTCCATGAGGCTCTCAGCC	540	
DB	481	CATAGTCCCTGCTGAGCCTATTATATACACCTGCTCGGCTCCATGAGGCTCTCAGCC	540	
QY	541	CAAAACCTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCCATACATTTCAATC	600	
DB	541	CAAAACCTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCCATACATTTCAATC	600	
QY	601	CTGTTTCTGAAACAACTTTCAGCACAGAAATAACACCACTTCCGTTTGTAGTA	660	
DB	601	CTGTTTCTGAAACAACTTTCAGCACAGAAATAACACCACTTCCGTTTGTAGTA	660	
QY	661	GGACCTCTGTTTCCCAATCTGGAATAAACCATACCTCAAACTCTGTTGTAATAATT	720	
DB	661	GGACCTCTGTTTCCCAATCTGGAATAAACCATACCTCAAACTCTGTTGTAATAATT	720	
QY	721	AGCAATACTATAGACACACAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGA	780	
DB	721	AGCAATACTATAGACACACAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGA	780	
QY	781	ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG	840	
DB	781	ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG	840	
QY	841	AATGGCTCTTCAGAACTCATGCTCTCTCTCTTCTAGTGGCCCTTATGACCATCTAC	900	
DB	841	AATGGCTCTTCAGAACTCATGCTCTCTCTCTTCTAGTGGCCCTTATGACCATCTAC	900	
QY	901	ACTGAACAAAGATTTATACAATCATGTGCTAGCTAAGCCCGCAACAAGAGTACCCATT	960	
DB	901	ACTGAACAAAGATTTATACAATCATGTGCTAGCTAAGCCCGCAACAAGAGTACCCATT	960	

QY	961	CTTCCCTTTTGTATATCAGAGCAGAGTGCTAGGCAGACTAGGTACTGGCATTTGGCAGTATC	1020
DB	961	CTTCCCTTTTGTATATCAGAGCAGAGTGCTAGGCAGACTAGGTACTGGCATTTGGCAGTATC	1020
QY	1021	ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGAATGGAACAG	1080
DB	1021	ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGAATGGAACAG	1080
QY	1081	GTCACTGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT	1140
DB	1081	GTCACTGACTCCCTGGTCACTTTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT	1140
QY	1141	CAAAATCGAAGAGCTTTAGACTTTGTAACCGCCAAAGAGGGGGAACCTCTTTATTTTA	1200
DB	1141	CAAAATCGAAGAGCTTTAGACTTTGTAACCGCCAAAGAGGGGGAACCTCTTTATTTTA	1200
QY	1201	GGAGAAGACGCTGTTATTAATCAATCCAGAAATGCTCACTGAGAAAGTTAAAGAA	1260
DB	1201	GGAGAAGACGCTGTTATTAATCAATCCAGAAATGCTCACTGAGAAAGTTAAAGAA	1260
QY	1261	ATTGAGATCGAATACAAATAGAGCAGAGAGCTTCAAAACACCGACGCTGGGGCTC	1320
DB	1261	ATTGAGATCGAATACAAATAGAGCAGAGAGCTTCAAAACACCGACGCTGGGGCTC	1320
QY	1321	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
DB	1321	CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
QY	1381	TTACTCTCTTTGGACCTGTAATCTTTAACTCTCTTAACTTTGCTCTTCCAGAAAT	1440
DB	1381	TTACTCTCTTTGGACCTGTAATCTTTAACTCTCTTAACTTTGCTCTTCCAGAAAT	1440
QY	1441	GAACTGTAAAGCTACAGATGCTCTCAAAATGGAACCCCA	1481
DB	1441	GAACTGTAAAGCTACAGATGCTCTCAAAATGGAACCCCA	1481

RESULT 2	1481 bp	mRNA	linear	PAT 17-AUG-2003
AR344387	Sequence 105 from patent US 6582703.			
LOCUS	AR344387			
DEFINITION	AR344387			
ACCESSION	AR344387.1	GI:33740328		
VERSION				
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 1481)			
AUTHORS	Perron,H., Beseme,F., Bedin,F., Paranhos-Baccala,G., Komurian-Pradel,F., Jolivet-Reynaud,C. and Mandrand,B.			
TITLE	Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting			
JOURNAL	Patent: US 6582703-A 105 24-JUN-2003;			
FEATURES	Location/Qualifiers			
source	1..1481			
	/organism="unknown"			
	/mol_type="mRNA"			

ORIGIN	Query Match 100.0%; Score 1481; DB 6; Length 1481;		
	Best Local Similarity 100.0%; Pred. No. 0;		
	Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGGCCCTCCCTATCATACACTTTTCTCTTTACTGTTCTCTTACCCCTTTTGGCTCTCACT	60
DB	1	ATGGCCCTCCCTATCATACACTTTTCTCTTTACTGTTCTCTTACCCCTTTTGGCTCTCACT	60
QY	61	GCACCCCTCCATGCTGTTACACACGATAGCTCCCTTACCAAGAGTTTCTATGAAGA	120
DB	61	GCACCCCTCCATGCTGTTACACACGATAGCTCCCTTACCAAGAGTTTCTATGAAGA	120
QY	121	ACGCGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC	180
DB	121	ACGCGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC	180
QY	181	ACCTTCACCTGCCACACCCATATGCCCGCACTGCTATACTCTGCCACTCTTTGCATG	240
DB	181	ACCTTCACCTGCCACACCCATATGCCCGCACTGCTATACTCTGCCACTCTTTGCATG	240
QY	241	CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT	300
DB	241	CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTCTAGTTGCTCTGGAGACTT	300
QY	301	GGAGCCACTGCTGTTGGACTTACTTACCCATACAGTATGCTCTGATGGGGTGAATT	360
DB	301	GGAGCCACTGCTGTTGGACTTACTTACCCATACAGTATGCTCTGATGGGGTGAATT	360
QY	361	CAAGTCAGCAAGAGAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT	420
DB	361	CAAGTCAGCAAGAGAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT	420
QY	421	AGCACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAACTACATGAACCCCTCCGTACC	480
DB	421	AGCACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAACTACATGAACCCCTCCGTACC	480
QY	481	CATAGTCCCTGCTGAGCCTATTATACACCTGCTCGGCTCCATGAGGCTCTCAGCC	540
DB	481	CATAGTCCCTGCTGAGCCTATTATATACACCTGCTCGGCTCCATGAGGCTCTCAGCC	540
QY	541	CAAAACCTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCCATACATTTCAATC	600
DB	541	CAAAACCTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCCATACATTTCAATC	600
QY	601	CTGTTTCTGAAACAACTTTCAGCACAGAAATAACACCACTTCCGTTTGTAGTA	660
DB	601	CTGTTTCTGAAACAACTTTCAGCACAGAAATAACACCACTTCCGTTTGTAGTA	660
QY	661	GGACCTCTGTTTCCCAATCTGGAATAAACCATACCTCAAACTCTGTTGTAATAATT	720
DB	661	GGACCTCTGTTTCCCAATCTGGAATAAACCATACCTCAAACTCTGTTGTAATAATT	720
QY	721	AGCAATACTATAGACACACAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGA	780
DB	721	AGCAATACTATAGACACACAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGA	780
QY	781	ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG	840
DB	781	ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG	840
QY	841	AATGGCTCTTCAGAACTCATGCTCTCTCTCTTCTAGTGGCCCTTATGACCATCTAC	900
DB	841	AATGGCTCTTCAGAACTCATGCTCTCTCTCTTCTAGTGGCCCTTATGACCATCTAC	900
QY	901	ACTGAACAAAGATTTATACAATCATGTGCTAGCTAAGCCCGCAACAAGAGTACCCATT	960
DB	901	ACTGAACAAAGATTTATACAATCATGTGCTAGCTAAGCCCGCAACAAGAGTACCCATT	960


```
Db 361 CAAGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCGGTACC 480
Db 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCGGTACC 480
Qy 481 CATACTCGCTGGTGGCTATTTAATACCAACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
Db 481 CATACTCGCTGGTGGCTATTTAATACCAACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
Qy 541 CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACTTCAGGCCATACATTTCAATC 600
Db 541 CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACTTCAGGCCATACATTTCAATC 600
Qy 601 CTGTTCTCTGAAACAATGGAACAACCTTCAGCAGAGAAATAACACACATTCCTGTTTAGTA 660
Db 601 CTGTTCTCTGAAACAATGGAACAACCTTCAGCAGAGAAATAACACACATTCCTGTTTAGTA 660
Qy 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCTACTCAAACTCCTCACTGTCTAAATTT 720
Db 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCTACTCAAACTCCTCACTGTCTAAATTT 720
Qy 721 AGCAATACTATAGACAACAACCCAGCTCCCAATGATCAGGTGGGTAAACACTCCCAACAGA 780
Db 721 AGCAATACTATAGACAACAACCCAGCTCCCAATGATCAGGTGGGTAAACACTCCCAACAGA 780
Qy 781 ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGTGTTACCTCAGCCTATCATTTGTTG 840
Db 781 ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGTGTTACCTCAGCCTATCATTTGTTG 840
Qy 841 AATGGCTCTTCAGAACTATGCTTCCTCTCTATCTTAGTGCCCTCTAGCCACTCTAC 900
Db 841 AATGGCTCTTCAGAACTATGCTTCCTCTCTATCTTAGTGCCCTCTAGCCACTCTAC 900
Qy 901 ACTGAAACAAGATTTATACAACTCATGTCGTAACCTTAAGCCCCCAACAACAAAGAGTACCATT 960
Db 901 ACTGAAACAAGATTTATACAACTCATGTCGTAACCTTAAGCCCCCAACAACAAAGAGTACCATT 960
Qy 961 CTTCTCTTTGTTATCAGACAGAGGTGCTAGGACACTAGGTACTGGCAATGGCGATATC 1020
Db 961 CTTCTCTTTGTTATCAGACAGAGGTGCTAGGACACTAGGTACTGGCAATGGCGATATC 1020
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACATATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 1021 ACAACCTCTACTCAGTTCTACTACAACATATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Qy 1081 GTCACCTGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Db 1081 GTCACCTGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Qy 1141 CAAAATCGAAGAGCTTTAGACTTGCTTAACCGCCAAAGAGGGGGAACCTGTTTATTTTA 1200
Db 1141 CAAAATCGAAGAGCTTTAGACTTGCTTAACCGCCAAAGAGGGGGAACCTGTTTATTTTA 1200
Qy 1201 GGAGAGAACGCTGTTATTTATGTTTAACTCAGAAATGTCATCAGAAAGTAAAGAA 1260
Db 1201 GGAGAGAACGCTGTTATTTATGTTTAACTCAGAAATGTCATCAGAAAGTAAAGAA 1260
Qy 1261 ATTGAGATCGAATACAATAGAGCAGAGGAGCTTCAAAACACCGAAGCGCTGGGCGCTC 1320
Db 1261 ATTGAGATCGAATACAATAGAGCAGAGGAGCTTCAAAACACCGAAGCGCTGGGCGCTC 1320
Qy 1321 CTAGCCAAATGGATGCGCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAAATTTG 1380
Db 1321 CTAGCCAAATGGATGCGCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAAATTTG 1380
Qy 1381 TTACTCTCTTTGGACCCCTGATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGAAAT 1440
Db 1381 TTACTCTCTTTGGACCCCTGATCTTTAACTCTCTGTTAAAGTTTGTCTCTCCAGAAAT 1440
Qy 1441 GAAGCTGTAAGCTTACAGATGGTCTTTACAAATGGAACCCCA 1481
Db 1441 GAAGCTGTAAGCTTACAGATGGTCTTTACAAATGGAACCCCA 1481
```

RESULT 4

AF127228

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

misc_feature

misc_difference

ORIGIN

Query Match

Best Local Similarity

Matches 1481; Conservative

Score 100.0%;

Pred. No. 0;

Mismatches 0;

Indels 0;

Gaps 0;

Length 1932;

DB 14;

Query 1

ATGGCCCTCCCTTATCATACCTTTTCTCTTACTGTCTCTTACCCCTTTGCTCTCACT 60

Db 452 ATGGCCCTCCCTTATCATACCTTTTCTCTTACTGTCTCTTACCCCTTTGCTCTCACT 511

Qy 61 GCACCCCTCCATGCTGTGTACACCAAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 120

Db 512 GCACCCCTCCATGCTGTGTACACCAAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 571

Qy 121 AGCGGCTTCCTGGAAATATTGATGCCCATATATAGAGTTTATTAAGGGAACCTCC 180

Db 572 AGCGGCTTCCTGGAAATATTGATGCCCATATATAGAGTTTATTAAGGGAACCTCC 631

Qy 181 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Db 632 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 691

Qy 241 CATGCAAACTACATATTATGGACAGGAAATGATTAATCTTAGTTGCTCGGAGACTT 300

Db 692 CATGCAAACTACATATTATGGACAGGAAATGATTAATCTTAGTTGCTCGGAGACTT 751

Qy 301 GGAGCCACTGTCTGTTGGACTTACTTCCACCATACAGTATCTGTAGTGGGTGGAATT 360

Db 752 GGAGCCACTGTCTGTTGGACTTACTTCCACCATACAGTATCTGTAGTGGGTGGAATT 811

AF127228 1932 bp mRNA linear VRL 11-AUG-1999
Multiple sclerosis associated retrovirus element clone CL15
env-like mRNA sequence.

AF127228
AF127228.1 GI:5726293

Multiple sclerosis associated retrovirus element
Multiple sclerosis associated retrovirus element
Viruses; Retroviral viruses; Retroviridae.

1 (bases 1 to 1932)

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F.,
Ounanian-Paraz, A., Sodoier, M., Ott, C., Rajoharison, A., Garcia, E.,
Mallet, F., Mandrand, B. and Perron, H.
Molecular cloning and characterization of MSRV-related sequences
associated with retrovirus-like particles

Virology 260 (1), 1-9 (1999)

9935590

10405350

2 (bases 1 to 1932)

Komurian-Pradel, F., Paranhos-Baccala, G., Bedin, F.,
Ounanian-Paraz, A., Sodoier, M., Ott, C., Rajoharison, A., Garcia, E.,
Mallet, F., Mandrand, B. and Perron, H.

Direct Submission

Submitted (10-FEB-1999) UMR103 CNRS, bioMerieux, 46, allée

D'Italie, Lyon 69007, France

Location/Qualifiers

1. .1932

/organism="Multiple sclerosis associated retrovirus

element"

/mol_type="mRNA"

/db_xref="taxon:89382"

/clones="CL15"

452. .1932

/note="similar to env but interrupted by premature stop

codon; includes signal peptide region, and partial SU and

TM regions"

misc_difference 568

/note="in clones CL15-2 and CL37-14; results in a

premature stop codon to Trp amino acid translation"

/replace="g"

Query Match

Best Local Similarity

Matches 1481; Conservative

Score 100.0%;

Pred. No. 0;

Mismatches 0;

Indels 0;

Gaps 0;

Length 1932;

DB 14;

Query 1

ATGGCCCTCCCTTATCATACCTTTTCTCTTACTGTCTCTTACCCCTTTGCTCTCACT 60

Db 452 ATGGCCCTCCCTTATCATACCTTTTCTCTTACTGTCTCTTACCCCTTTGCTCTCACT 511

Qy 61 GCACCCCTCCATGCTGTGTACACCAAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 120

Db 512 GCACCCCTCCATGCTGTGTACACCAAGTAGTCCCTTACCAAGAGTTCTTATGAAGA 571

Qy 121 AGCGGCTTCCTGGAAATATTGATGCCCATATATAGAGTTTATTAAGGGAACCTCC 180

Db 572 AGCGGCTTCCTGGAAATATTGATGCCCATATATAGAGTTTATTAAGGGAACCTCC 631

Qy 181 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240

Db 632 ACCTTCACTGCCCACACCATATGCTCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 691

Qy 241 CATGCAAACTACATATTATGGACAGGAAATGATTAATCTTAGTTGCTCGGAGACTT 300

Db 692 CATGCAAACTACATATTATGGACAGGAAATGATTAATCTTAGTTGCTCGGAGACTT 751

Qy 301 GGAGCCACTGTCTGTTGGACTTACTTCCACCATACAGTATCTGTAGTGGGTGGAATT 360

Db 752 GGAGCCACTGTCTGTTGGACTTACTTCCACCATACAGTATCTGTAGTGGGTGGAATT 811

repeat_region	1143. .1456 /note="AluXs repeat: matches 1. .312 of consensus"	misc_feature	16067. .16294 /note="match: GSS: Em:AQ054794 Em:AQ823559"
repeat_region	1457. .2938 /note="L1M8C repeat: matches 356. .1897 of consensus"	repeat_region	16923. .16998 /note="match: GSS: Em:AQ457003"
repeat_region	2922. .3739 /note="L1M2A repeat: matches 5329. .6146 of consensus"	repeat_region	17107. .17166 /note="L2 repeat: matches 2688. .2750 of consensus"
repeat_region	3844. .4087 /note="L1M4 repeat: matches 1. .251 of consensus"	repeat_region	17175. .17755 /note="L2 repeat: matches 1244. .1898 of consensus"
repeat_region	4092. .4231 /note="L1M4 repeat: matches 612. .756 of consensus"	repeat_region	17839. .18106 /note="MER70A repeat: matches 148. .418 of consensus"
repeat_region	4373. .4559 /note="L1M8C repeat: matches 971. .1152 of consensus"	repeat_region	18213. .18318 /note="L2 repeat: matches 1128. .1247 of consensus"
repeat_region	4586. .4619 /note="L1M2A repeat: matches 1. .313 of consensus"	repeat_region	18706. .18811 /note="L2 repeat: matches 2631. .2750 of consensus"
repeat_region	4623. .4964 /note="L1M2A repeat: matches 3. .333 of consensus"	repeat_region	18812. .19057 /note="L2 repeat: matches 1. .297 of consensus"
repeat_region	4989. .5149 /note="L1M8 repeat: matches 6120. .6288 of consensus"	repeat_region	19058. .19492 /note="L2 repeat: matches 2213. .2647 of consensus"
repeat_region	5152. .5602 /note="L1M2A repeat: matches 5325. .5779 of consensus"	misc_feature	complement(19352. .19843) /note="match: GSS: Em:AQ340313"
repeat_region	5606. .5916 /note="L1M8C repeat: matches 1. .313 of consensus"	repeat_region	19556. .19872 /note="L2 repeat: matches 1. .326 of consensus"
repeat_region	5917. .6014 /note="L1M2A repeat: matches 3. .333 of consensus"	repeat_region	19873. .21594 /note="L2 repeat: matches 6812. .8523 of consensus"
repeat_region	6018. .6409 /note="L1M2A repeat: matches 5769. .6157 of consensus"	repeat_region	21587. .22280 /note="L2 repeat: matches 4301. .4980 of consensus"
repeat_region	6566. .6750 /note="L1M4 repeat: matches 2857. .3036 of consensus"	repeat_region	22298. .22742 /note="L1M8C repeat: matches 5623. .6135 of consensus"
repeat_region	6751. .7063 /note="L1M4 repeat: matches 1. .307 of consensus"	repeat_region	22743. .23048 /note="L1M8C repeat: matches 1. .306 of consensus"
repeat_region	7064. .8097 /note="L1M4 repeat: matches 3036. .4105 of consensus"	repeat_region	23049. .23286 /note="L1M8C repeat: matches 5403. .5623 of consensus"
repeat_region	8099. .8419 /note="L1M8C repeat: matches 1. .306 of consensus"	repeat_region	23285. .23361 /note="L1M4 repeat: matches 5006. .5082 of consensus"
repeat_region	8420. .8487 /note="L1M8C repeat: matches 5582. .5648 of consensus"	repeat_region	23364. .23429 /note="L1M4 repeat: matches 4019. .4084 of consensus"
repeat_region	10039. .10396 /note="L1M8C repeat: matches 1. .364 of consensus"	repeat_region	23429. .24305 /note="L1 repeat: matches 3097. .4043 of consensus"
repeat_region	10519. .10807 /note="L1M8C repeat: matches 3. .297 of consensus"	repeat_region	24306. .24593 /note="L1M8C repeat: matches 1. .293 of consensus"
repeat_region	11343. .11416 /note="L1M8C repeat: matches 1. .297 of consensus"	repeat_region	24594. .25109 /note="L1 repeat: matches 2600. .3097 of consensus"
repeat_region	11448. .11983 /note="L1M8C repeat: matches 7162. .7739 of consensus"	repeat_region	25110. .25436 /note="L1M8C repeat: matches 5101. .5440 of consensus"
repeat_region	12030. .12337 /note="L1M8C repeat: matches 1. .309 of consensus"	repeat_region	25437. .25742 /note="L1M8C repeat: matches 1. .307 of consensus"
repeat_region	12995. .13077 /note="L1M8C repeat: matches 103. .186 of consensus"	repeat_region	25779. .25961 /note="L1M8C repeat: matches 55. .260 of consensus"
repeat_region	13110. .13222 /note="L1M8C repeat: matches 2. .124 of consensus"	repeat_region	25962. .26059 /note="L1M8C repeat: matches 1. .89 of consensus"
repeat_region	13228. .13518 /note="L1M8C repeat: matches 2131. .2414 of consensus"	repeat_region	26135. .26328 /note="L1M8C repeat: matches 144. .352 of consensus"
repeat_region	13519. .13853 /note="L1M8C repeat: matches 1. .346 of consensus"	repeat_region	26554. .26697 /note="L1M8C repeat: matches 1. .149 of consensus"
repeat_region	13854. .14039 /note="L1M8C repeat: matches 1951. .2131 of consensus"	repeat_region	26711. .26936 /note="L1M8C repeat: matches 5939. .6162 of consensus"
repeat_region	14040. .14254 /note="L1M8C repeat: matches 1. .218 of consensus"	repeat_region	26981. .27154 /note="L1M8C repeat: matches 5620. .5805 of consensus"
repeat_region	14255. .14807 /note="L1M8C repeat: matches 1686. .1951 of consensus"	repeat_region	27182. .27307 /note="L1M8C repeat: matches 5454. .5584 of consensus"
repeat_region	14808. .15105 /note="L1M8C repeat: matches 1. .311 of consensus"		
repeat_region	15106. .15492 /note="L1M8C repeat: matches 1309. .1686 of consensus"		
repeat_region	15496. .15822 /note="L1M8C repeat: matches 5823. .6167 of consensus"		
repeat_region	15824. .15920 /note="L1M8C repeat: matches 1044. .1140 of consensus"		
misc_feature	15885. .16086 /note="match: GSS: Em:AQ769824"		
misc_feature	15920. .16082 /note="match: GSS: Em:AQ769824"		
misc_feature	complement(15927. .16073) /note="match: GSS: Em:AQ769824"		

QY 121 ACGGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 180
DB 21439 ACGGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 21380
QY 181 ACCTTCACCTGCCACACCCATATATGCGCGCAACTGCTATACTCTGCGCACTCTTTGCATG 240
DB 21379 ACCTTCACCTGCCACACCCATATATGCGCGCAACTGCTATACTCTGCGCACTCTTTGCATG 21320
QY 241 CATGCAAACTACTATTTGGACAGGGAATGATTAATCTTAGTTGTCTGGAGACTT 300
DB 21319 CATGCAAACTACTATTTGGACAGGGAATGATTAATCTTAGTTGTCTGGAGACTT 21260
QY 301 GGAGCCACTGCTGTTGGACTTACTTACCCATACCACTATGTCATGATGGGGTGAAT 360
DB 21259 GGAGCCACTGCTGTTGGACTTACTTACCCATACCACTATGTCATGATGGGGTGAAT 21200
QY 361 CRAAGTCAGGCAAGAGAAAACAAAGTAAAGGAAGCAATCTCCAACTGACCGGGGACAT 420
DB 21199 CRAAGTCAGGCAAGAGAAAACAAAGTAAAGGAAGCAATCTCCAACTGACCGGGGACAT 21140
QY 421 AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
DB 21139 AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 21080
QY 481 CATACTCGCTGGTGGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCTTCAGCC 540
DB 21079 CATACTCGCTGGTGGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCTTCAGCC 21020
QY 541 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTTCAGGCGCATACATTTCAATC 600
DB 21019 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTTCAGGCGCATACATTTCAATC 20960
QY 601 CTTGTTCTTGAA CAATGGAAACAATTCTCAGACAGAAATAAACAACCTTCCTGTTTAGTA 660
DB 20959 CTTGTTCTTGAA CAATGGAAACAATTCTCAGACAGAAATAAACAACCTTCCTGTTTAGTA 20900
QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCATCTCAAACTCTCACTCTGTGTAATAATTT 720
DB 20899 GGACCTCTGTTTCCAACTCTGGAATAAACCCATCTCAAACTCTCACTCTGTGTAATAATTT 20840
QY 721 AGCAATACTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCCAACGA 780
DB 20839 AGCAATACTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCCAACGA 20780
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 840
DB 20779 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 20720
QY 841 AATGGCTCTTCAGAACTATGCTCTCTCATTTCTTAGTGCCCTATGACCATCTAC 900
DB 20719 AATGGCTCTTCAGAACTATGCTCTCTCATTTCTTAGTGCCCTATGACCATCTAC 20660
QY 901 ACTGAAACAAGATTTATACAATCATGTGCTACCTAAGCCCAACCAAAAGAGTACCCATT 960
DB 20659 ACTGAAACAAGATTTATACAATCATGTGCTACCTAAGCCCAACCAAAAGAGTACCCATT 20600
QY 961 CTTTCTTTTGTATCAGACAGGAGTCTAGGAGACTAGGTACTGGCATGGCAGTATC 1020
DB 20599 CTTTCTTTTGTATCAGACAGGAGTCTAGGAGACTAGGTACTGGCATGGCAGTATC 20540
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTGACATGGACAG 1080
DB 20539 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTGACATGGACAG 20480
QY 1081 GTCACTGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCT 1140
DB 20479 GTCACTGACTCCCTGGTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCT 20420
QY 1141 CAAATCGAAGAGCTTTAGACTGTCTAACCGCCAAAAGAGGGGAACTCTGTTTATTTTAA 1200
DB 20419 CAAATCGAAGAGCTTTAGACTGTCTAACCGCCAAAAGAGGGGAACTCTGTTTATTTTAA 20360

QY 1201 GGAGAGAAACGCTGTATTATGTTTAAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
DB 20359 GGAGAGAAATGCTGTATTATGTTTAAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 20300
QY 1261 ATTCAGATCGAATCAATAGTAGACAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 1320
DB 20299 ATTCAGATCGAATCAATAGTAGACAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 20240
QY 1321 CTCAGCAATGATGATCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 20239 CTCAGCAATGATGATCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG 20180
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTAACTCTCTTTAAAGTTGTCTCTTCCAGAAAT 1440
DB 20179 TTACTCTCTTTGGACCCCTGTATCTTAACTCTCTTTAAAGTTGTCTCTTCCAGAAAT 20120
QY 1441 GAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAAACCCCA 1481
DB 20119 GAAGCTGTAAAGCTACAAATGGTTCTTCAATGGAGCCCA 20079
RESULT 6
QY19381
LOCUS Sequence 5315 from Patent WO02068579. linear PAT 03-FEB-2004
DEFINITION QY19381
ACCESSION QY19381
VERSION QY19381.1 GI:42280238
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter C.J., Adams M.C., Li P.W. and Myers E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 5315 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Source Location/Qualifiers
1..2055
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.0%; Score 1466.6; DB 6; Length 2055;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1472; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATACTTTCTTTACTTGTCTCTTACCCCTTTTCGCTCTCACT 60
DB 36 ATGGCCCTCCCTTATCATATTTTCTTTACTTGTCTTCTTACCCCTTTTCGCTCTCACT 95
QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
DB 96 GCACCCCTCCATGCTGTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 155
QY 121 ACGGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 180
DB 156 ACGGCGCTTCTGGAATATTGATGCCCATATAGGAGTTTATCTAAGGAACTCC 215
QY 181 ACCTTCACCTGCCACACCCATATGCGCGCAACTGTCTATAACTTGCACCTCTTTGCATG 240
DB 216 ACCTTCACCTGCCACACCCATATGCGCGCAACTGTCTATACTTGCACCTCTTTGCATG 275
QY 241 CATGCAAACTACTATTTGGACAGGGAATGATTAATCTTAGTTGTCTCTGGAGACTT 300
DB 276 CATGCAAACTACTATTTGGACAGGGAATGATTAATCTTAGTTGTCTCTGGAGACTT 335
QY 301 GGAGCCACTGCTGTTGGAGCTTACTTACCCATACCACTATGTCATGATGGGGTGAAT 360
DB 336 GGAGCCACTGCTGTTGGAGCTTACTTACCCATACCACTATGTCATGATGGGGTGAAT 395

QY 361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
Db 396 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 455
QY 421 AGACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC 480
Db 456 AGCACCCTTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC 515
QY 481 CATACTCGCTCGTGAAGCTATTTAATACCACTTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 516 CATACTCGCTCGTGAAGCTATTTAATACCACTTCACTCGGCTCCATGAGGTCTCAGCC 575
QY 541 CAAAACCTTACTAATCTGTTGGATGTGCTCCCTCGCATTCAGGCCATACATTTCAATC 600
Db 576 CAAAACCTTACTAATCTGTTGGATGTGCTCCCTCGCATTCAGGCCATACATTTCAATC 635
QY 601 CTGTTCTGAACTAGGAACACTTTCAGACAGAAATAAACCACTTCCGTTTTAGTA 660
Db 636 CTGTTCTGAACTAGGAACACTTTCAGACAGAAATAAACCACTTCCGTTTTAGTA 695
QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCTCACTGTGTAAATTT 720
Db 696 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCTCACTGTGTAAATTT 755
QY 721 AGCAATCTATAGACACAACCACTCCCAATGATCAGTGGGTAAACCTCCACACGA 780
Db 756 AGCAATCTATAGACACAACCACTCCCAATGATCAGTGGGTAAACCTCCACACGA 815
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCATCATGTTGTTG 840
Db 816 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCATCATGTTGTTG 875
QY 841 AATGGCTCTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCTTAGACCACTATC 900
Db 876 AATGGCTCTTCAGAACTATGTGCTTCTCTCAATCTTAGTGCCCTTAGACCACTATC 935
QY 901 ACTGAACAAGTTTATACAACTATGTGCTACCTTAAGCCCCACAAAGAGTACCATT 960
Db 936 ACTGAACAAGTTTATACAACTATGTGCTACCTTAAGCCCCACAAAGAGTACCATT 995
QY 961 CTTCTCTTTGTTATCAGACGAGGTGCTAGGACAGACTAGGTACTGGCATTTGCCATATC 1020
Db 996 CTTCTCTTTGTTATCAAGCAGGNGTCTAGGACAGACTAGGTACTGGCATTTGCCATATC 1055
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAGAAATAAATGGTGAATGGAACAG 1080
Db 1056 ACAACCTCTACTCAGTTCTACTACAACTATCTCAGAAATAAATGGTGAATGGAACAG 1115
QY 1081 GTCACTGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGAGTAGTCCCT 1140
Db 1116 GTCACTGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGAGTAGTCCCT 1175
QY 1141 CAAAATCGAAGAGCTTTAGACTTCTTAACCGCCCAAAAAGAGGGGAACTCTTTATTTTA 1200
Db 1176 CAAAATCGAAGAGCTTTAGACTTCTTAACCGCCCAAAAAGAGGGGAACTCTTTATTTTA 1235
QY 1201 GGAGAAGAACGCTGTTATATGTTAATCAATCCAGAAATGTCATGAGAAGTTAAGAA 1260
Db 1236 GGAGAAGAAATGCTGTTATATGTTAATCAATCCAGAAATGTCATGAGAAGTTAAGAA 1295
QY 1261 ATTCGAGATCGAATACAAATGAGAGAGAGGCTTCAAAACACCGAACGCTGGGGCTC 1320
Db 1296 ATTCGAGATCGAATACAAATGAGAGAGAGGCTTCAAAACACCGAACGCTGGGGCTC 1355
QY 1321 CTCAGCCAATGGATGCCCTGGGTTCTCCCTTTCTAGGAAGCTCTTAAATTTG 1380
Db 1356 CTCAGCCAATGGATGCCCTGGGTTCTCCCTTTCTAGGAAGCTCTTAAATTTG 1415
QY 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAGTTGTTCTCTCCAGAAAT 1440
Db 1416 TTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAGTTGTTCTCTCCAGAAAT 1475

QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGAACCCCA 1481
Db 1476 GAAGCTGTAAGCTACAAATGGTCTTCAATGGAGCCCA 1516
RESULT 7
LOCUS BD252040 1629 bp DNA linear PAT 17-JUL-2003
DEFINITION Method for detecting superantigen activity in a biological sample.
ACCESSION BD252040
VERSION BD252040.1 GI:33061810
KEYWORDS JP 2002539804-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1629)
AUTHORS Perron, H. and Lafont, M.
TITLE Method for detecting superantigen activity in a biological sample
JOURNAL Patent: JP 2002539804-A 1 26-NOV-2002;
BIO MERIEUX
COMMENT OS Homo sapiens (human)
PN JP 2002539804-A/1
PD 26-NOV-2002
PF 20-MAR-2000 JP 2000607009 99/13755 PI
PR 19-MAR-1999 FR 99/03622, 28-OCT-1999 FR
HERVE PERRON, MONIQUE LAFONT
PC C12N15/09, A61K31/711, A61K38/00, A61K39/395, A61K45/00, A61K48/00,
PC A61P25/00,
PC C07K14/47, C07K16/18, C12N7/00, C12Q1/02, C12Q1/68, G01N27/447, PC
G01N33/15,
PC
G01N33/50, G01N33/569, G01N33/577//C12P21/08, (C12N7/00, C12R1:93), PC
(C12Q1/02, C12R1:93), (C12Q1/68, C12R1:93), C12N15/00, G01N27/26, PC
A61K37/02
CC Method for detecting superantigen activity in a biological CC
FH Key Location/Qualifiers
FT source 1..1629 /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..1629 /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.7%; Score 1461.8; DB 6; Length 1629;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1469; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATACCTTTCTTTACTGTCTTCTTACCCCTTTGCTCTCACT 60
Db 1 ATGGCCCTCCCTTATCATACCTTTCTTTACTGTCTTCTTACCCCTTTGCTCTCACT 60
QY 61 GCACCCCTCCATGCTGTGTACACCACTAGTCTCCCTTACCAAGAGTTCTTATGAAGA 120
Db 61 GCACCCCTCCATGCTGTGTACACCACTAGTCTCCCTTACCAAGAGTTCTTATGAAGA 120
QY 121 AGCGGCTTCCCTGGAATAATTCATGCCCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 121 AGCGGCTTCCCTGGAATAATTCATGCCCATATAGGAGTTTATCTAAGGGAACCTCC 180
QY 181 ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATACTTCGCCACTCTTTGCATG 240
Db 181 ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATACTTCGCCACTCTTTGCATG 240
QY 241 CATGCAAACTACTATTTGGACAGGAAATGATTAATCCTAGTTGTCCTGGAGACTT 300
Db 241 CATGCAAACTACTATTTGGACAGGAAATGATTAATCCTAGTTGTCCTGGAGACTT 300
QY 301 GGAGCCACTGTCTGTTGGAGTTTACTTCAACCCATACAGATATGTCGATGGGGTGAAT 360

Qy 121 ACGGGCTTCTCGGAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180
Db 103354 ACGGGCTTCTCGGAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAACTCC 103413
Qy 181 ACTTTCACTGCCACACCCATATGCCCCGGAACCTGCTATTAACCTCGCCACTCTTTGCAATG 240
Db 103414 ACTTTCACTGCCACACCCATATGCCCCGGAACCTGCTATTAACCTCGCCACTCTTTGCAATG 103473
Qy 241 CATGCAAACTACTCATATTGGAAGGAAATGATTAATCTAGTTCTCTGCGAGGACTT 300
Db 103474 CATGCAAACTACTCATATTGGAAGGAAATGATTAATCTAGTTCTCTGCGAGGACTT 103533
Qy 301 GGAGCCACTGTCTGTGAGCTTACTTCAACCCATACCAAGTATGCTGATGCGGGGTTGAATT 360
Db 103534 GGAGCCACTGTCTGTGAGCTTACTTCAACCCATACCAAGTATGCTGATGCGGGGTTGAATT 103593
Qy 361 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGACAT 420
Db 103594 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGACAT 103653
Qy 421 AGCACCCCTAGCCCTCAAGAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db 103654 AGCACCCCTAGCCCTCAAGAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 103713
Qy 481 CATACTCGCCCTGTGAGCTTATTAAATACCAACCTCACTCGGCTCCATGAGTCTCAGCC 540
Db 103714 CATACTCGCCCTGTGAGCTTATTAAATACCAACCTCACTCGGCTCCATGAGTCTCAGCC 103773
Qy 541 CAAACCCCTACTAATCTGTGATGCTGCTCCCTCGCACTTTCAGGCCATATCATTTCAATC 600
Db 103774 CAAACCCCTACTAATCTGTGATGCTGCTCCCTCGCACTTTCAGGCCATATCATTTCAATC 103833
Qy 601 CCTGTTCTGGAACATGGAACAATCTCAGCACAGAAATAACACCACTTCGCTTTTATGTA 660
Db 103834 CCTGTTCTGGAACATGGAACAATCTCAGCACAGAAATAACACCACTTCGCTTTTATGTA 103893
Qy 661 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAACCTCAGCTGTGTAATAATTT 720
Db 103894 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAACCTCAGCTGTGTAATAATTT 103953
Qy 721 AGCAATACTATPAGACACAACAGCTCCCAATGCAATCAGTGGGTAAACCTCCACACGA 780
Db 103954 AGCAATACTATPAGACACAACAGCTCCCAATGCAATCAGTGGGTAAACCTCCACACGA 104013
Qy 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGCTGCTAGCTCAGCTATCATTTGTTG 840
Db 104014 ATAGTCTGCTACCTCAGGAATATTTTGTCTGCTGCTAGCTCAGCTATCATTTGTTG 104073
Qy 841 AATGGCTCTCAGAACTATGCTGCTCTCTCATTTAGTGGCCCTATGACCATCTAC 900
Db 104074 AATGGCTCTCAGAACTATGCTGCTCTCTCATTTAGTGGCCCTATGACCATCTAC 104133
Qy 901 ACTGAACAAGATTTATCAATCATGTGCTACCTAAGCCCCCAACAAAGAGTACCCATT 960
Db 104134 ACTGAACAAGATTTATCAATCATGTGCTACCTAAGCCCCCAACAAAGAGTACCCATT 104193
Qy 961 CTTCTTTTGTATCAGACGAGGAGTCTAGGAGACTAGTACTGCGATTGGCAGTATC 1020
Db 104194 CTTCTTTTGTATCAGACGAGGAGTCTAGGAGACTAGTACTGCGATTGGCAGTATC 104253
Qy 1021 ACAACCTCTACTAGTTTCTACTACAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
Db 104254 ACAACCTCTACTAGTTTCTACTACAACTATCTCAAGAAATAAATGTTGACATGGAACAG 104313
Qy 1081 GTCACCTGACTCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGAGTATGCTCTT 1140
Db 104314 GTCACCTGACTCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGAGTATGCTCTT 104373
Qy 1141 CAAATCGAAGAGCTTTAGACTTGTACCCGCAAGAGGGGAGACCTGTTTATTTTA 1200
Db 104374 CAAATCGAAGAGCTTTAGACTTGTACCCGCAAGAGGGGAGACCTGTTTATTTTA 104433
Qy 1201 GGAGAAGAACGCTGTTATTATGTTAAATCAATCCAGAAATTTGCTCACTGAGAAAGTTAAAGAA 1260

Db 104434 GGAGAAGAACGCTGTTATTATTATGTTAAATCAATCAGAAATTTGCTCACTGAGAAAGTTAAAGAA 104493
Qy 1261 ATTCGAGATCGAATACAAATGATAGAGAGAGCTTCAAAAACACCGAACGCTGGGGCTC 1320
Db 104494 ATTCGAGATCGAATACAAATGATAGAGAGAGCTTCAAAAACACCGAACGCTGGGGCTC 104553
Qy 1321 CTGAGCAATGGATGCCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 104554 CTGAGCAATGGATGCCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATTG 104613
Qy 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGAAAT 1440
Db 104614 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGAAAT 104673
Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481
Db 104674 GAAGCTGTAAGCTACAAATGGTCTTCAAAATGGAGCCCCA 104714

RESULT 11
AX478550
LOCUS AX478550 2074 bp DNA linear PAT 12-AUG-2002
DEFINITION Sequence 4 from Patent: WO0248362.
ACCESSION AX478550
VERSION AX478550.1 GI:22217319
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Ramkumar, J. and Arvizu, C.
TITLE Embryogenesis associated proteins
JOURNAL Patent: WO 0248362-A 4 20-JUN-2002;
INCYTE Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1.2074
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7477736CB1"

ORIGIN
Query Match 93.8%; Score 1389.8; DB 6; Length 2074;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1424; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTTCTTACCCCTTTGCGTCTCACT 60
Db 64 ATGGCCCTCCCTTATGTTATTTTCTTCTTACTGTTCTTACCACCTTTTCTCACT 123
Qy 61 GCACCCCTCCATGCTGCTGTACACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 120
Db 124 GCACCCCTCCATGCTGCTGTACACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 183
Qy 121 ACGCGGCTTCTCGAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180
Db 184 ATGGCGGCTTCCAGACATATTGATGCCCCCATCTGATAGGAGTTTATCTAAGGGAACTCC 243
Qy 181 ACCTTCACTGCCACACCCATATGCCCCGGAACCTGCTATTAACCTGCGACTTTTTCATG 240
Db 244 GCCTTCAACGCCACACCCATATGCCCCCACTGCTATAACTCTGCGACTTTTGTATG 303
Qy 241 CATGCAAACTACTCATATTGGAAGGAAATGATTAATCTAGTTCTCTGCGAGGACTT 300
Db 304 CATGCAAACTACTCATATTGGAAGGAAATGATTAATCTAGTTCTCTGCGAGGACTT 363
Qy 301 GGAGCCACTGTCTGTGAGCTTACTTCAACCCATACCAAGTATGCTGATGCGGGTGAATT 360
Db 364 GGAGCCACTGTCTGTGAGCTTACTTCAACCCATACCAAGTATGCTGATGCGGGTGAATT 423
Qy 361 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGACAT 420

QY 61 GCACCCCTCCATGCTGCTGTAACAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 7882 GCACCCCTCCATGCTGCTGTAACAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 7941
QY 121 ACSCGGCTCTCTGGAATATTTGATGCCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 7942 ATGCGGCTCTCTGGAATATTTGATGCCCCATCGATAGGAGTTTCTAAGGGAACCTCC 8001
QY 181 ACCTTCACTGCCACACCCATATGCCCCGCACTGCTATACCTCTGCTGCTCTTTGCTATG 240
Db 8002 ACCTTCACTGCCACACCCATATGCCCCGCACTGCTATACCTCTGCTGCTCTTTGCTATG 8061
QY 241 CATGCAAACTACTCATTTATGAGCAGGGAATGATTAATCCTAGTTCTCTGAGGAGCTT 300
Db 8062 CATGCAAACTACTCATTTATGAGCAGGGAATGATTAATCCTAGTTCTCTGAGGAGCTT 8121
QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCAAGTATGCTGATGGGGTGGAAATT 360
Db 8122 GGAGTCACTGTCTGTGGACTTACTTCAACCATACCAAGTATGCTGATGGGGTGGAGTT 8181
QY 361 CAAGGTCAGCAAGAGAAAACAAGTAAGGAGCAATCTCCCACTGACCCCGGGACAT 420
Db 8182 CAAGATCAGCAAGAGAAAACAAGTAAGGAGCAATCTCCCACTGACCCCGGGTACAT 8241
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db 8242 AGCACCTTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 8301
QY 481 CATACTCGCCTGTGAGCCTATTTAATPACCAACCTCTACTCGGCTCCATGAGGCTTCAGCC 540
Db 8302 CATACTCGCCTGTGAGCCTATTTAATPACCAACCTCTACTCGGCTCCATGAGGCTTCGCC 8361
QY 541 CAACACCTACTAATCTGTGATGCTGCTCCCTGCTCACTTCAGGCCATATCTCAATC 600
Db 8362 CAACACCTACTAATCTGTGATGCTGCTCCCTGCTCACTTCAGGCCATATCTCAATC 8421
QY 601 CCTGCTCTGAACTGAACTCTCAGACAGAAATAACACACTTCGTTTAACTAGTA 660
Db 8422 CCTGCTCTGAACTGAACTCTCAGACAGAAATAACACACTTCGTTTAACTAGTA 8481
QY 661 GGACCTCTGTTTCCAATCTGGAATAAACCCATACCTCAACCTCACTGTGTAATTT 720
Db 8482 GGACCTCTGTTTCCAATCTGGAATAAACCCATACCTCAACCTCACTGTGTAATTT 8541
QY 721 AGCAATCTATAGACACACACAGCTCCCAATGCATCAGTGGGTAAACCTCCACAGGA 780
Db 8542 AGCAATCTATAGACACACACAGCTCCCAATGCATCAGTGGGTAAACCTCCACACAA 8601
QY 781 ATAGTCTGCTTACCTCAGGAATATTTTCTGCTGCTACTCTCAGGCTATCATTTGTTG 840
Db 8602 ATAGTCTGCTTACCTCAGGAATATTTTCTGCTGCTACTCTCAGGCTATCATTTGTTG 8661
QY 841 AATGGCTCTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 8662 AATGGCTCTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8721
QY 901 ACTGAACAAATTTATACATCATCTGTGTACTTAAGCCCCCAACAAGAGTAGTACCATT 960
Db 8722 ACTGAACAAATTTATACATCATCTGTGTACTTAAGCCCCCAACAAGAGTAGTACCATT 8781
QY 961 CTTCCTTTTGTATCAGACGAGGAGTGTAGGACACTTAGTACTTGGCATTGGCAGTATC 1020
Db 8782 CTTCCTTTTGTATAGGACGAGGAGTGTAGGACACTTAGTACTTGGCATTGGGCTATC 8841
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 8842 ACNACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGGACATGGAACAG 8901
QY 1081 GTCATGATCTCCCTGGTCACTTGGCAAGATCAATTAATCTCCCTAGCAGGAGTAGTCTT 1140
Db 8902 GTCGCCGACTCCCTGGTCACTTGGCAAGATCAATTAATCTCCCTAGCAGGAGTAGTCTT 8961
QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAAACCTGTTTATTTT 1200

Db 8962 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGAAACCTGTTTATTTT 9021
QY 1201 GGAGAGAGAGCGCTGTTTATTTATTTATCAATCAGAAATTTGCTCACTGAGAAAGTTAAAGAA 1260
Db 9022 GGAGAGAGAGCGCTGTTTATTTATTTATCAATCAGAAATCGTCACTGAGAAAGTTAAAGAA 9081
QY 1261 ATTGAGAGTCGAATCAATATGAGAGCAGAGGAGCTTCAAAACACCGACGCTGGGCGCTC 1320
Db 9082 ATTGAGAGTCGAATCAACCTAGAGCAGAGGAGCTTCAAAACACCGACGCTGGGCGCTC 9141
QY 1321 CTCAGCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db 9142 CTCAGCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9201
QY 1381 TTACTCTCTCTTTGGACCTGCTATCTTTAACTCTCTTTAAAGTTTGTCTCTCCAGAAAT 1440
Db 9202 CTACTCTCTTTGGACCTGCTATCTTTAACTCTCTTTAACTCTCTTTAACTCTCTCTCTCCAGAAAT 9261
QY 1441 GAAGCTGTAAAGCTACAGATGCTTTCAAAATGGAACCCCA 1481
Db 9262 GAAGCTGTAAACTCAAAATGGAGCCCAAGATGAGTCCAA 9302

RESULT 13

AC145964
LOCUS
DEFINITION
AC145964
VERSION
AC145964.3 GI:36016769
KEYWORDS
HTG.
SOURCE
Pan troglodytes (chimpanzee)

184675 bp DNA linear PRI 29-OCT-2003
Pan troglodytes BAC clone RP43-12F2 from 7, complete sequence.

ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Shahid, S., Cotton, M., Bielicki, L. and Meyer, R.
The sequence of Pan troglodytes BAC clone RP43-12F2
Unpublished (2001)

2 (bases 1 to 184675)
Sulston, J.E. and Wilson, R.
Sequencing of Pan troglodytes
Unpublished (2001)

3 (bases 1 to 184675)
Wilson, R.K.
Direct Submission
Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

4 (bases 1 to 184675)
Wilson, R.K.
Direct Submission
Submitted (24-SEP-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

5 (bases 1 to 184675)
Wilson, R.K.
Direct Submission
Submitted (26-SEP-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA

6 (bases 1 to 184675)
Wilson, R.
Direct Submission
Submitted (29-OCT-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Sep 25, 2003 this sequence version replaced gi:35073486.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu>
Contact: submissions@wustl.edu

Summary Statistics
Center project name: C_PT012F02

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA was isolated from white blood cells obtained from a male chimpanzee (Pan troglodytes, 'Clint', Verkes #C0471; birthdate: 6-6-80). The clone and detailed information can be obtained from ResGen (<http://www.resgen.com>) or Pieter de Jong and co-workers at <http://www.bacpac.chori.org>.

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES		Location/Qualifiers	
source	1..184675	/organism="Pan troglodytes"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9598"	
		/chromosome="7"	
		/map="7"	
		/clone="RP43-12P2"	
		/clone_lib="RPCI-43"	
unsure	115419..115500	/note="Sequence derived from one plasmid subclone."	
unsure	138694..138701	/note="Sequence derived from one plasmid subclone."	
misc_feature	140277..140388	/note="Sequence derived from PCR product of project DNA."	
ORIGIN			
Query Match 90.6%; Score 1341.8; DB 9; Length 184675;			
Best Local Similarity 94.1%; Pred. No. 0;			
Matches 1394; Conservative 0; Mismatches 87; Indels 0; Gaps 0;			
Qy	1	ATGGCCCTCCCTTATCATATTTTCTCTTACTGTCTTTACCCCTTTTCACTCTCACT	60
Db	153623	ATGGCCCTCCCTTATCATATTTTCTCTTACTGTCTTTTACCCCTTTTCACTCTCACT	153682
Qy	61	GCACCCCTCCATGCTGTGTACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA	120
Db	153683	GCACCCCTCCATGCTGTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGGAGA	153742
Qy	121	ACGGGGCTTCTGGAAATATTGATGCCCGCCCATATAGGAGTTTATCTAAGGGAATCC	180
Db	153743	ATGCAGCGTCCCGGAAATATTGATGCCCGCCCATATAGAGATTTTCTAAGGGAACCCCC	153802
Qy	181	ACCTTCACTGCCACACCCATATGCCCGGCAACTGCTATAACTCTGCCACTTTTGGCATG	240
Db	153803	ACCTTCACTGCCACACCCATATGCCCGGCAACTGCTATCACTCTGCCACTTTTGGCATG	153862
Qy	241	CATGCAATATCTATTATTGACAGGGAATGATTAATCCTAGTTGTCTCGAGGACTT	300
Db	153863	CATGCAATATCTATTATTGACAGGGAATGATTAATCCTAGTTGTCTCGAGGACTT	153922
Qy	301	GGAGCCACTGTCTGTGGACTTACTTACCCATACCAAGTATGCTGTATGGGGGTGGAAAT	360
Db	153923	GGAGTCACTGTCTGTGGACTTACTTACCCATACCAAGTATGCTGTATGGGGGTGGAGTT	153982

Qy	361	CAAGGTCAGCAAGAGAAAAAAGTAAGAAAGCAATCTCCCAACTGACCCGGGACAT	420
Db	153983	CAAGATCAGCAAGAGAAAAAAGTAAGAAAGCAATCTCCCAACTCACCAGGATACAT	154042
Qy	421	AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTCATGATAAACCCTCCGTACC	480
Db	154043	AGCACCTTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTCATGATAAACCCTCCGTACC	154102
Qy	481	CATACCTCGCTGTGAGCCCTATTATACCAACCTCAGCTCGGCTCCATGAGGCTCAGCC	540
Db	154103	CATACCTCGCTGTGAGCCCTATTATACCAACCTCAGCTCGGCTCCATGAGGCTCAGCC	154162
Qy	541	CAAAACCTCTACTACTGTGGATGTGCTCCCTCCCTGCACTTCAGGCCATACATTTCAATC	600
Db	154163	CAAAACCTCTACTACTGTGGATGTGCTCCCTCCCTGCACTTCAGGCCATATGTTTCAATC	154222
Qy	601	CCTGTTCTGTGAACAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTATGA	660
Db	154223	CCTGTTCTGTGAACAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTATGA	154282
Qy	661	GGACCTCTGTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCAGCTGTGTAATTTT	720
Db	154283	GGACCTCTGTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCAGCTGTGTAATTTT	154342
Qy	721	AGCAATCTATAGACACAAACCACTCCCAATGCAATGCAATGCAATGCAATGCAATGCA	780
Db	154343	AGCAATCTATAGACACAAACCACTCCCAATGCAATGCAATGCAATGCAATGCAATGCA	154402
Qy	781	ATAGTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGTGCTAGCTCAGCCATCATGTTTG	840
Db	154403	ATAGTCTGCCCTACCCCTCAGGAATATTTTGTCTGTGTGCTAGCTCAGCCATCATGTTTG	154462
Qy	841	AATGCTCTTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900
Db	154463	AATGCTCTTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	154522
Qy	901	ACTGAAACAAGATTATACAAATCATGCTGCTAAGCCCAACCAAAAGAGTACCCATT	960
Db	154523	ACTGAAACAAGATTATACAAATCATGCTAAGCCCAACCAAAAGAGTACCCATT	154582
Qy	961	CTTCTCTTTTGTATCAGACGAGGAGTGTAGGAGACTAGTACTGCGCATTTGGCAGTATC	1020
Db	154583	CTTCTCTTTTGTATCAGACGAGGAGTGTAGGAGACTAGTACTGCGCATTTGGCAGTATC	154642
Qy	1021	ACAACTCTACTCAGTCTTCTATACAACTATCTCAAGAAATAAATGGTGCATGGAACAG	1080
Db	154643	ACAACTCTACTCAGTCTTCTATACAACTATCTCAAGAAATAAATGGTGCATGGAACAG	154702
Qy	1081	GTCACTGACTCCCTGGTCACTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT	1140
Db	154703	GTCCGCGACTCCCTGGTCACTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT	154762
Qy	1141	CAAAATCGAAGAGCTTTAGACTGCTTAACCGCAAGAGGGGGAACCTGTTTATTTT	1200
Db	154763	CAAAATCGAAGAGCTTTAGACTGCTTAACCGCAAGAGGGGGAACCTGTTTATTTT	154822
Qy	1201	GGAGGAAGACGCTGTTTATTTATGTAATCAATCCAGAAATTTGTCACTGAGAAAGTTAA	1260
Db	154823	GGAGGAAGATGCTGTTTATTTATGTAATCAATCCAGAAATTTGTCACTGAGAAAGTTAA	154882
Qy	1261	ATTTCAGAGATCGAATCAATGTTAGACAGAGAGCTTCAAAACACCGAAGCTGGGGCTTC	1320
Db	154883	ATTTCAGAGATCGAATCAATGTTAGACAGAGAGCTTCAAAACACCGAAGCTGGGGCTTC	154942
Qy	1321	CTCAGCAATGAGTCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG	1380
Db	154943	CTCAGCAATGAGTCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG	155002
Qy	1381	TTACTCCTCTTTGGAGCCCTGTATCTTTAACTCTTTTAACTTTTAACTTTTGTCTCTCC	1440
Db	155003	CTACTCCTCTTTGGAGCCCTGTATCTTTAACTCTTTTAACTTTTAACTTTTGTCTCTCC	155062

QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGAACTGTTATTTTAA 1200
Db CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGAACTGTTATTTTAA 9018
QY 1201 GGAGAGAAACGCTGTTATTTATTTATCAATCCAGAAATGTGACATGAGAAAGTTAAAGAA 1260
Db CCGGAGAAATGCTGTTATTTATTTATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 9078
QY 1261 ATTCGAGATCGAATCAATGTTAGAGCAGAGGAGCTTCAAAAACCCGAAACGCTGGGGCCTC 1320
Db ATTCGAGATCGAATCAATGTTAGAGCAGAGGAGCTTCAAAAACCCGAAACGCTGGGGCCTC 9138
QY 1321 CTCAGCAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATG 1380
Db CTCAGCAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATG 9198
QY 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAACTTTGTTCTCTTCCAGAAAT 1440
Db CTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAACTTTGTTCTCTTCCAGAAATC 9258
QY 1441 GAAGCTGTAAGCTACAGATGCTTACAAATGGAACCCCA 1481
Db GAAGCTGTAAGCTACAAATGAGCCCAAGATGCAGTCCAA 9299

RESULT 15
AY101587 10229 bp DNA linear PRI 11-FEB-2004
LOCUS Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVW1
DEFINITION locus, allele B, complete sequence.
ACCESSION AY101587
VERSION AY101587.1 GI:37544409
KEYWORDS Pan troglodytes (chimpanzee)
SOURCE Pan troglodytes
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 10229)
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G.,
Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.
The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 10229)
Mallet, F., Bouton, O. and Oriol, G.
Direct Submission
TITLE Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-BioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..10229
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/isolate="1"
/isolation_source="PBMC of chimpanzee 1"
/db_xref="taxon:9598"
/cell_type="PBMC"
/note="isolated by PCR; ERVW1 locus, allele B
endogenous_virus: ERV-W"
1..779
join(248..810,2889..9820)
/note="putative mRNA transcript 1; gag-pol pseudogene"
join(248..810,7592..9820)
/product="envelope glycoprotein"
/note="putative mRNA transcript 2"
join(248..810,9255..9820)
/note="putative mRNA transcript 3; no coding region"
7819..9435
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"

/protein_id="AAM68166.1"
/db_xref="GI:37544410"
/translation="MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYQBFLLRMORP
GNIDAPYSRFSKGTFTTAHTHMPNCHVSHATLCKHANTHRYPYKGLDLSKLHETLR
TVCTWTFYTHGMSDGGVQDQAREKHVKEISQLTRVHSASRYPKGLDLSKLHETLR
HTLVLSTFNTLTGLHEVSAQNTNCMLPLNFRPYVSPVPEQWNFNSTEINTSV
LVGLVSLNLEITHTSNLTCVKFSNTTDTTNSQCIRWVPTPTQILPSPGFFVCGGTA
YRCINGSSSEMCFLSFLVPMTIYTBQDLYNVVYVPRKRVPLIFVIGAGVGAAG
TGIGGTTSTQFYFKLSOELNGDMERVADSLTLODLQNSLAALVQNRRLDILLTAE
RGCTCLFLGEECCYYVNSGIVTEKVEIRLRIORBAEELRTNGPGLLSQMWPEWLP
FLGLAAIILLLLFGPCIFNLVNFVSSRIEAVLQNEPKMQSKTKIYRPLDRPASP
RSDVNDIKGTTPPEILTAOPLLRPNAGS"
9494..10229
LTR
ORIGIN
Query Match 90.5%; Score 1340.2; DB 9; Length 10229;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTATCATACTTTTCTTTACTGTCTCTTACCCCTTTGGCTCTCACT 60
Db ATGGCCCTCCCTATCATATTTTCTTTACTGTCTTTTACCCTCTTTCACCTCACT 7878
QY 61 GCACCCCTCCATGCTGTFACAAACAGTAGTCCCTTACCAAGATTTCTATGAAGA 120
Db GCACCCCTCCATGCCGCTGTATGCCAGTAGTCCCTTACCAAGATTTCTATGAAGA 7938
QY 121 ACGGGCTTCTGGAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180
Db ATGCAGCGTCCCGAAATATTTGATGCCCATCATATAGGAGTTTCTAAGGGAACTCC 7998
QY 181 ACCTTCATCTGCCACACCCCATATGCCCGCAACTGTCTATAACTCTGCCACTCTTTGCA 240
Db ACCTTCATCTGCCACACCCCATATGCCCGCAACTGTCTATAACTCTGCCACTCTTTGCA 8058
QY 241 CATGCAATATCTATTTATTTGACAGGGAATATTAATCTAGTTCTCTGGAGGACTT 300
Db CATGCAATATCTATTTATTTGACAGGGAATATTAATCTAGTTCTCTGGAGGACTT 8118
QY 301 GGAGCACTGCTGTGGACTTCTTACCCCATACCACTAGTATGTCTGATGGGGTGGAAAT 360
Db GGAGTCACTGCTGTGGACTTCTTACCCCATACTGTTATGTCTGATGGGGTGGAGTT 8178
QY 361 CAAAGTCAGGCAAGAGAAAAAACAAGTAAAGAAAGCAATCTCCCACTGACCCGGGACAT 420
Db CAAAGTCAGGCAAGAGAAAAAACAAGTAAAGAAAGCAATCTCCCACTGACCCGGGACAT 8238
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC 480
Db AGCGCCTCTAGCCCTACAAAGGACTAGTCTCTCAAACTACATGAACCTCCGTACC 8298
QY 481 CATACTCGCTGTGAGGCTATTTAATACCACTCTCACTCGGCTCCATGAGTCTCAGCC 540
Db CATACTCGCTGTGAGGCTATTTAATACCACTCTCACTCGGCTCCATGAGTCTCAGCC 8358
QY 541 CAAAAACCTTACTAATGTTGGATGCTCCCTCGCATCTTCCAGCCCATACATTTCAATC 600
Db CAAAAACCTTACTAATGTTGGATGCTCCCTCGCATCTTCCAGCCCATACATTTCAATC 8418
QY 601 CTGTTCTGAACTGGAACCAACTTACAGCAGAGAAATAAACCACTTCCGTTTTAGTA 660
Db CTGTTCTGAACTGGAACCAACTTACAGCAGAGAAATAAACCACTTCCGTTTTAGTA 8478
QY 661 GGACCTCTGTTTCCAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 720
Db GGACCTCTGTTTCCAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 8538
QY 721 AGCAATCTATAGACAAACAGCTCCCAATGCATCAGTGGTGAACCTCCACACCA 780
Db AGCAATCTATAGACAAACAGCTCCCAATGCATCAGTGGTGAACCTCCACACCA 8598
QY 781 ATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
Db ATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG

Db 8599 ATAGTCTGCTACCCCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTG 8658

Qy 841 AATGGCTCTTCAAACTATGCTCTTCTCTATTTCTTAGTGCCTTATGACCATCTAC 900

Db 8659 AATGGCTCTTCAAACTATGCTCTTCTCTATTTCTTAGTGCCTTATGACCATCTAC 8718

Qy 901 ACTGAACAAGATTATATCAATCATGCTGATCTTAAAGCCCAACAAAGAGTACCCATT 960

Db 8719 ACTGAACAAGATTATATCAATCATGCTGATCTTAAAGCCCAACAAAGAGTACCCATT 8778

Qy 961 CTTCCCTTTTATACAGAGCAGAGTCTAGGAGACTAGGTACTGGCATGGCAGTATC 1020

Db 8779 CTTCCCTTTTATAGGAGCAGAGTCTAGGTGCACTAGGTACTGGCATGGCAGTATC 8838

Qy 1021 ACAACCTCTACTCAGTCTTACTCAAACTATCTCAAGAAATTAATGTGACATGGAACAG 1080

Db 8839 ACAACCTCTACTCAGTCTTACTCAAACTATCTCAAGAAATTAATGTGGAACAGG 8898

Qy 1081 GTCACCTGACTCCCTGCTGCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT 1140

Db 8899 GTCGCCGACTCCCTGCTGCTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTT 8958

Qy 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGAGGGGAACTGTTTATTTTA 1200

Db 8959 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTTA 9018

Qy 1201 GGAGAAGACGCTTTATTTATTTATTAATCAATCAGATTTCTCACTGAGAAGTTAAGAA 1260

Db 9019 GGGGAAGAACTGTTATTTATTTATTAATCAATCAGATTTCTCACTGAGAAGTTAAGAA 9078

Qy 1261 ATTGAGATCGAATACAACTGTAGAGCAGAGAGCTTCAAAACACCGAAGCTGGGGCTC 1320

Db 9079 ATTGAGATCGAATACAACTGTAGAGCAGAGAGCTTCAAAACACCGAAGCTGGGGCTC 9138

Qy 1321 CTCAGCAATGAGTGGCTTGGTCTTCCCTTCTTAGGACTCTAGCAGTCTTAATTTG 1380

Db 9139 CTCAGCAATGAGTGGCTTGGTCTTCCCTTCTTAGGACTCTAGCAGTCTTAATTTG 9198

Qy 1381 TTACTCTCTTTGGACCTGTATCTTAACTCCCTGCTTAAAGTTGCTCTTCCAGAAAT 1440

Db 9199 CTACTCTCTTTGGACCTGTATCTTAACTCCCTGCTTAAAGTTGCTCTTCCAGAAAT 9258

Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481

Db 9259 GAAGCTGTAAGCTACAAATGGAGCCCAAGATGAGTCCCA 9299

RESULT 16

LOCUS AV101588 10230 bp DNA linear PRI 11-FEB-2004

DEFINITION Gorilla gorilla isolate 1 endogenous retrovirus ERV-W, ERVW1 locus, allele A, complete sequence.

ACCESSION AV101588

VERSION AV101588.1 GI:37544411

KEYWORDS

SOURCE Gorilla gorilla (gorilla)

ORGANISM Gorilla gorilla

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.

AUTHORS Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L., and Mandrand, B.

TITLE The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

PUBMED 14757826

REFERENCE 2 (bases 1 to 10230)

AUTHORS Mallet, F., Bouton, O., and Oriol, G.

TITLE Direct Submission

JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France

FEATURES Location/Qualifiers

1..10230

/organism="Gorilla gorilla"

/mol_type="genomic DNA"

/isolate="1"

/isolation_source="PBMC of gorilla 1"

/db_xref="taxon:9593"

/cell_type="PBMC"

/note="isolated by PCR; ERVW1 locus, allele A endogenous_virus: ERV-W"

1..795

join(248..826,2894..9821)

/note="putative mRNA transcript 1; gag-pol pseudogene"

join(248..826,7595..9821)

/note="putative mRNA transcript 2"

/product="envelope glycoprotein"

join(248..826,9258..9821)

/note="putative mRNA transcript 3; no coding region"

7822..9438

/note="syncytin"

/codon_start=1

/product="envelope glycoprotein"

/protein_id="AAM68167.1"

/db_xref="GI:37544412"

/translation="MALPYHLLFTVLLPSFTLTAPPCCMTSSSPYOEFLWRMRP GNIDAPSHRSFSGKGTFTFATHMPRNCNSATLCMHANTHYWTGMNPSCPGGLV TVCTVYFTHTGMSDGGVQDQAREKHVEVTSQITRVHSTSSPYKGLDLSKHEITRT HTRLVSLFNTLTGLHEVAQNPTNCWICLPDIFRYSIVPVEWNFSKEHTSTV LVGLVSNLEIHTSNLTNLTQKFSNTIDTNSOCIRWVTPPTQVCLPISGIPFVCGTSA YLGLSSSSSMFLSLPVPMTIYTDQILVNVVSKPRKRVPIPLFVIGAGVLGALG TGLGTTSTFYQYKLSQBLNGDMERVAADSLVTLQDLNSLAAVLONRALDOLTAE RGCTCLFGECCYVYVNSQI VTEKVEIRDRIRORABELRNTGPNGLLSQWNPWLP FLGPLAAIILLLLFGFCIFNLVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP RSDVNDIKCTPPEEISTAQPLLRLPNSGSS"

9497..10230

LTR

ORIGIN

Query Match 90.5%; Score 1340.2; DB 9; Length 10230;

Best Local Similarity 94.1%; Pred. No. 0;

Matches 1393; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATATCTTCTTCTTACGTGTTCTTACCCCTTTCCGCTCTCACT 60

Db 7822 ATGGCCCTCCCTTATCATATCTTCTTCTTACGTGTTCTTACCCCTTTCCGCTCTCACT 7881

Qy 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120

Db 7882 GCACCCCTCCATGCTGTGTATGACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 7941

Qy 121 ACGGCGCTTCTGGAATATTGATGCCCATATATAGGAGTTTATCTAAGGGAATCTC 180

Db 7942 ATGGCGCTTCTGGAATATTGATGCCCATATGATGAGGAGTTTCTTAAAGGGAATCTC 8001

Qy 181 ACCTTACTGCCACACCCATATGCCCGCACTGTATTAAGTCTGCTCACTCTTTGATG 240

Db 8002 ACCTTACTGCCACACCCATATGCCCGCACTGTATTAAGTCTGCTCACTCTTTGATG 8061

Qy 241 CATGCAATACTCATTTATGACAGGAAATGATTAATCTAGTCTTCTTGGAGGACTT 300

Db 8062 CATGCAATACTCATTTATGACAGGAAATGATTAATCTAGTCTTCTTGGAGGACTT 8121

Qy 301 GGAGCCACTGCTGTTGGACTTACTTCAACCATACAGTATGTCTGATGGGGTGGAAAT 360

Db 8122 GGAGTCACTGCTGTTGGACTTACTTCAACCATACAGTATGTCTGATGGGGTGGAGTT 8181

Qy 361 CAAGGTGAGCAAGAGAAAACAAGTAAAGAGCAATCTCCCACTGACCCGGGGACAT 420

Db 8182 CAAGATCAGGCAAGAGAAAACAAGTAAAGAGCAATCTCCCACTGACCCGGGGTACAT 8241

Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAATACATGATAAACCCTCCGTACC 480

Db 8242 AGCACCTTAGCCCTACAAAGGACTAGATCTCTCAAAATACATGATAAACCCTCCGTACC 8301

Qy 481 CATACTCGCTGTGGTGGCTTATTTAATACCAACCTCACTCGGTCCATGAGGTCTCAGCC 540


```
QY 181 ACCTTCACCTGCCACACCATATGCCCGCAACTGCTATAAATCTGCGCACTCTTTGCATG 240
Db 7986 ACCTTCACCTGCCACACCATATATGCGCGCACTGCTATAAATCTGCGCACTCTTTGCATG 8045
QY 241 CATGCAAAATCTCATTTATGGAAGGAAATGATTAATCTAGTTGCTCTGAGGACTTT 300
Db 8046 CATGCAAAATCTCAGTATTTGGAAGGAAATGATTAATCTAGTTGCTCTGAGGACTTT 8105
QY 301 GGAGCCACTGCTGTTGGACTTACTTACCCATACAGATATGTCATGATGGGGTGAAT 360
Db 8106 GGAGTCACTGCTGTTGGACTTACTTACCCATACAGATATGTCATGATGGGGTGAAT 8165
QY 361 CAAAGTTCAGGCAAGCAAGAAACAAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
Db 8166 CAAAGTTCAGGCAAGCAAGAAACAAAGTAAAGGAAGTAACTCTCCAACTGACCCAGGTACAT 8225
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAATCTACATGAAGACCTTCCGTACC 480
Db 8226 AGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAATCTACATGAAGACCTTCCGTACC 8285
QY 481 CATACTCGCTGCTGAGCCTTATTAATACACCTCAGTCCGCTCCATGAGGTCTCAGCC 540
Db 8286 CACACTCGCTGCTGAGCCTTATTAATACACACTCAGTCCGCTCCATGAGGTCTCGGCC 8345
QY 541 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
Db 8346 CAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACTTCAGGCGCATATGTTTCAATC 8405
QY 601 CTTGTTCTGTAACAATGGAACAACTTCAGCACAGAAATAAACAACCACTTCCTCGTTTGTAGTA 660
Db 8406 CTTGTTCTGTAACAATGGAACAACTTCAGCACAGAAATAAACAACCACTTCCTCGTTTGTAGTA 8465
QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATTT 720
Db 8466 GGACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATTT 8525
QY 721 AGCAATACTATAGACACACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACAGA 780
Db 8526 AGCAATACTATAGACACACAGCTCCCAATGATCAGGTGGGTAACTCTCTCCACACAGA 8585
QY 781 ATAGTCTGCTACCCCTCAGGAATATTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 840
Db 8586 ATATTCGCTACCCCTCAGGAATATTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 8645
QY 841 AATGGCTCTCAGAACTATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 8646 AATGGCTCTCAGAACTATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8705
QY 901 ACTGAACAAGATTTATCAATCATGTCGTACCTAAGCCCAACCAAGAGTACCCATT 960
Db 8706 ACTGAACAAGATTTATCAATCATGTCGTATCTAAGCCCAACCAAGAGTACCCATT 8765
QY 961 CTTCTCTTTTGTATCAGACGAGGTGCTAGGACACTAGGTACTGCGCATTTGGCAGTATC 1020
Db 8766 CTTCTCTTTTGTATGAGGAGCAGGAGTCTAGGTGCTAGGTACTGCGCATTTGGCAGTATC 8825
QY 1021 ACAACCTCTACTGTTCTACTCAAACTATCTCAAGAAATTAATGTTGATGATGAGACAG 1080
Db 8826 ACAACCTCTACTGTTCTACTCAAACTATCTCGAGAATTAATGTTGATGATGAGACAG 8885
QY 1081 GTCACTGACTCCCTGGTCACTGTCGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCTTT 1140
Db 8886 GTCCGCCACTCCCTGGTCACTGTCGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCTTT 8945
QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAAGAGGGGAAACCTGTTTATTTT 1200
Db 8946 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAAACCTGTTTATTTT 9005
QY 1201 GGAGAGAAAGCTGTTTATTTATTTATTAATCAATCCAGAAATTTGCTACTGAGAAAGTTAAGAA 1260
Db 9006 GGGAGAAAGTCTGTTTATTTATTTATTTATTAATCAATCCGGAATTCGTTACTGAGAAAGTTAAGAA 9065
QY 1261 ATTCGAGATCGAATACATGTTAGACGAGGAGCTTCAAAACACCGAAGCTGGGGCCTC 1320
```

```
Db 9066 ATTGAGATCGAATACAACTAGACGAGGAGCTTCGAAACATTTGGACCCCTGGGGCCTC 9125
QY 1321 CTCACCCAAATGATCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 9126 CTCACCCAAATGATCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTAATATTG 9185
QY 1381 TTACTCTCTCTTTGGACCCCTGTAATTTAACTCTCTTTAAAGTTTGTCTCTCCAGAAAT 1440
Db 9186 CTGCTCTCTTTGGACCCCTGTAATTTAACTCTCTTTAACTCTCTCTCTCTCTCCAGAAAT 9245
QY 1441 GAAGCTGTAAAGCTACAGATGCTTACAAATGGAACCCCA 1481
Db 9246 GAAGCTATAAACTACAAATGAGCCCAAGATGGAGTCCCA 9286

RESULT 18
AY101592 10248 bp DNA linear PRI 11-FEB-2004
Hylobates pileatus isolate 1 endogenous retrovirus ERV-W, ERVWE1
locus, allele A, complete sequence.
ACCESSION AY101592
VERSION AY101592.1 GI:37544419
KEYWORDS
SOURCE Hylobates pileatus (pileated gibbon)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE 1 (bases 1 to 10248)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1..10248
/organism="Hylobates pileatus"
/mol_type="genomic DNA"
/isolate="1"
/isolation_source="PBMC of gibbon 1"
/db_xref="taxon:9589"
/cell_type="PBMC"
/notes="isolated by PCR; ERVWE1 locus, allele A
endogenous_virus: ERV-W"
1..773
join(248..804,2997..9809)
/notes="putative mRNA transcript 1; gag-pol pseudogene"
join(248..804,7581..9809)
/product="envelope glycoprotein"
/notes="putative mRNA transcript 2"
join(248..804,9244..9809)
/notes="putative mRNA transcript 3; no coding region"
7808..9424
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68171.1"
/db_xref="GI:37544420"
/translacion="MALPIHIFLFTVLLPSFTLTAPPCCRTSSSPQEFPLARTORP
GNIDAPLYRSFKSGPTFTAHYRKVEKVIQSQTQVHSTSSPYKGLDLSKLHETLRT
TVCWYFTHTGMSDGGVQDQAREKHVEKVIQSQTQVHSTSSPYKGLDLSKLHETLRT
HRLVSLFNLTTLGLHEVQONPCWMLPLDFRPYVSIPEQWNFNSTENITTSV
LVLGPLVNLIEITHNSLTCKFSTNTDNTSOCIWVTPTRIFCLPSGFFVCGTSA
YRCINGSSSMCFSLFVPMWIIYEQDLNYYVVKPRKRVPIIPVVMGAVLGALG
TGTGSIITTSQFYIKLSRLNGDMERVADSLVTLQDOLSLAAVILQNRRLDLITAE
RGCTCLFLGECCYYNQSGIVTEKVKIRDRIOORAEELRNIGPWLGLSQMWPWILP
```



```

misc_feature
43..57
/notes="Caucasian
endogenous_virus: HERV-W"
/notes="splice acceptor site"
284..1900
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17561.1"
/db_xref="GI:33410925"
/translation="MALPYHIFLTVLLPSFTLAPPPCRMTSSSPYQEFLEWRMORP
GNVADPSYSLSGSTFTFAHMPRENCYHSLCMHANTHYWTGMINPSPGGLGV
TICWTYFTQSGDGVODQAREKHVKEIVSOLTRVHGTSYKGLDLSKLHETLT
HYELVSLFTTLTGLHEVAQNPTNCWICLPLNFRPVSIPIVPEOWNFSEINTTSV
LGPLVSNLEIHTSHSLTCVKFSNTYITNSQCIRWTPPTQIVCLPSPGIFVCGTSA
YKLVNLSSEMSFLSPVPMWITITQDLYNYVISKPRNKRPFILFVIGAGVLGALG
TGIGITITSTQFYKLSOBLNGDMERVADSLVLODLSLAALVNLONRRALLDTAE
RGGTCLFLGEBECYVQSGIVTEKVEIRDRIOREBELRNTGPMGLSLSOWMPWLP
FLGPLAAIILLLLFGPCIFNLVNFVSSRIEAVKLQWEPKMQSKTKIYRPLDRPASP
RSDVNDIKGTPPEISAQPLLRLPNSAGSS"
1705..1720
/notes="splice acceptor site"
1959..2694

```

```

ORIGIN
Query Match          90.2%; Score 1335.4; DB 9; Length 2694;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGCGCCCTCTTATCATCTTTCTCTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60
DB 284 ATGCGCCCTCTTATCATATTTTCTTACTGTTCTTTTACCCCTTTTCACTCTCACT 343
QY 61 GCACCCCTCCATCTGCTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 344 GCACCCCTCCATGCGCTGTATGACCACTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 403
QY 121 ACGGGCTCTCTGGAATATGATGCCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB 404 ATGACGCTCTCCGGAATATATGATGCCCCATCTGTATAGGAGTTTCTTCTAAGGGAACCTCC 463
QY 181 ACCTTCACTGCCACACCATATGCCCCGCAACTGTCTATAACTCTGCCACTCTTTTGCATG 240
DB 464 ACCTTCACTGCCACACCATATGCCCCGCAACTGTCTATAACTCTGCCACTCTTTTGCATG 523
QY 241 CATGCAAAATCTCATTTATGACAGGAAATGATTAATCTAGTTCTCTGAGGAGCTT 300
DB 524 CATGCAAAATCTCATTTATGACAGGAAATGATTAATCTAGTTCTCTGAGGAGCTT 583
QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCATGATGCTGATGGGGTGAATT 360
DB 584 GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGCTGATGGGGTGGAGTT 643
QY 361 CAAGGTGAGGCAAGAGAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGGACAT 420
DB 644 CAAGATGAGGCAAGAGAAAACAATGTAAGAGAGTAACTCTCCCAACTCACCCGGGTACAT 703
QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAATPACATGAAGAACCTCCGTACC 480
DB 704 GGCACCTCTAGCCCTCAAAAGGACTAGTCTCTCAAAATPACATGAAGAACCTCCGTACC 763
QY 481 CATATCTGCGCTGGTGGAGCTTATTTAATACACCCCTCACTGGCTCCATGAGGTCTCAGCC 540
DB 764 CATATCTGCGCTGGTGGAGCTTATTTAATACACCCCTCACTGGCTCCATGAGGTCTCAGCC 823
QY 541 CAAACCCCTACTAATGTTGATGTGCTCCCTCGCACTTCAGGCCATATCAATTCAATC 600
DB 824 CAAACCCCTACTAATGTTGATGTGCTCCCTCGCACTTCAGGCCATATGTTTCAATC 883
QY 601 CCTGTCTCTGAACAAATGGAACAACTTACGACAGAAATAAACAACACTTCGTTTGTAGTA 660
DB 884 CCTGTACCTGAACAAATGGAACAACTTACGACAGAAATAAACAACACTTCGTTTGTAGTA 943

```

```

QY 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCTATACCTCAAACTCCTGCTGTAAATTT 720
DB 944 GGACCTCTTGTTCCTCAATCTGGAATAAACCCTATACCTCAAACTCCTGCTGTAAATTT 1003
QY 721 AGCAATATCTATAGACACACACACAGCTCCCAATCATCAGCTGGGTAAACCTCCACACAG 780
DB 1004 AGCAATATCTATAGACACACACACAGCTCCCAATCATCAGCTGGGTAAACCTCCACACAG 1063
QY 781 ATAGTCTGCTACTACCTCTCAGGAATAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
DB 1064 ATAGTCTGCTACTACCTCTCAGGAATAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1123
QY 841 AATGGCTCTTCAAGATCTATGCTCTCTCTCAATCTTCTTGTAGTGGCCCTATGACCATCTAC 900
DB 1124 AATGGCTCTTCAAGATCTATGCTCTCTCTCAATCTTCTTGTAGTGGCCCTATGACCATCTAC 1183
QY 901 ACTGAACAGATTTATACATCATCTGCTGCTTAAAGCCCAACAAAGAGAGTACCCATT 960
DB 1184 ACTGAACAGATTTATACATCATCTGCTGCTTAAAGCCCAACAAAGAGAGTACCCATT 1243
QY 961 CTTCTCTTTTGTATCAGACGAGGAGTCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020
DB 1244 CTTCTCTTTTGTATCAGACGAGGAGTCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1303
QY 1021 ACAACCTCTACTCAGTCTTACTACAAACTATCTCAAGAAATAATGGTGTGATGGAAACAG 1080
DB 1304 ACAACCTCTACTCAGTCTTACTACAAACTATCTCAAGAAATAATGGTGTGATGGAAACAG 1363
QY 1081 GTCACTGACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1364 GTCCGCGACTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGGAACCTGTTTATTTTAA 1200
DB 1424 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGGAACCTGTTTATTTTAA 1483
QY 1201 GGAAGAAACGCTGTTTATTTATGTTAATCAATCAAGAAATGCTCACTGAGAAAGTTAAAGAA 1260
DB 1484 GGAAGAAACGCTGTTTATTTATGTTAATCAATCAAGAAATGCTCACTGAGAAAGTTAAAGAA 1543
QY 1261 ATTGAGATCGAATACAATGATAGACAGAGGAGCTTCAAAACCGCAAGAGCTGGGGCTC 1320
DB 1544 ATTGAGATCGAATACAATGATAGACAGAGGAGCTTCAAAACCGCAAGAGCTGGGGCTC 1603
QY 1321 CTCAGCAATGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1380
DB 1604 CTCAGCAATGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1663
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 1440
DB 1664 CTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 1723
QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAAACCCCA 1481
DB 1724 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAAACCCCA 1764

```

```

RESULT 20
AF520487S2
LOCUS
DEFINITION
Homo sapiens individual 148 allele B, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION
AF520488
VERSION
AF520488.1 GI:33410944
KEYWORDS
2 of 2
SEGMENT
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2694)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.

```

```

AF520487S2 2694 bp DNA linear PRI 11-FEB-2004
Homo sapiens individual 148 allele B, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
AF520488
AF520488.1 GI:33410944
2 of 2
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2694)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.

```


TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology
JOURNAL Proc Natl Acad Sci U.S.A. 101 (6), 1731-1736 (2004)

PUBMED 14/5/826
 REFERENCE 2 (bases 1 to 2694)
 AUTHORS Mallet, F., Bouton, O. and Oriol, G.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
 CNRS-bioMérieux, Ecole Normale Supérieure de Lyon - 46 allée
 d'Italie, Lyon 69364 cedex 07, France

```

FEATURES             Location/Qualifiers
     source            ..2694
         /organism="Homo sapiens"
         /mol_type="genomic DNA"
         /isolate="individual 148 allele B"
         /db_xref="taxon:9606"
         /chromosome="7"
         /map="7q21-q22"
         /sex="male"
         /cell_type="PBMC's"
         /note="Caucasian"
         endogenous_virus: HERV-W"
misc_feature          43..57
                     /note="splice acceptor site"
CDS                   284..1900
                     /codon_start=1
                     /product="envelope glycoprotein"
                     /protein_id="AAQ17566.1"
                     /db_xref="GI:33410945"
                     /translation="MALPHIIFLFTLLPTAPPSCRCMTSSSPYQEPIRMQRPR
GNIDAPSYSLSKGTFTHATHPMKNVCYSATLCMHANTHWTKMGSSPGGLGV
HTRLVSLFDTITGLHEVSQAQDAREKHKEVEISQLTRVHGTSYPKGLDLSKLHLTAR
LVGPLVSNLEITHNLNLTCKVFSTTYTQDLNYNVISKPNRKPVPILPFVIGAGVLGLTA
YKLGNSSESMCLFLGPLPMTITYEODLYNVITSKPNRKVPILPFVIGAGVLGLTA
TGIGTTQTQPYKYLQBLNGDMERVAIVTLQDLNSLAADVLRQNRRALDILTAE
RGCTCLFLGECCYVNGSGITEVKIEIRDIRIQRAEELNTGPWGLLSOMFWPILF
FGPLAAAILLLLFGPCINLVNFVSRRIEAVKLQMEPKMQSKTKIYYRRLDRPASFI
RSDVDNIKGTPPEEISAQPLLRPNSAGSS"
                     1705..1720
                     /note="splice acceptor site"
                     1959..2694
                     LTR

```

```

LTR      1959.  .2694
ORIGIN
Query Match      90.2%; Score 1335.4; DB 9; Length 2694;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1200. Conservative 0. Mismatches 91; Indels 0; Gaps 0

```

[illegible]

QY	361	CAAGCTCAGGCAAGAGAAAAA	CAAGTAGTAAAGAAAGCAATCTCCCAACTGACCCGGGACAT	420
DB	644	CAAGATCAGGCAAGAGAAAAA	CATGTAAAGAGTAACTCCCAACTCACCCGGTACAT	703
QY	421	AGCACCCCTAGCCCTACAAAG	CACTAGTTCTCTCAAACTCATGAAACCCCTCCGTACC	480
DB	704	GGCACTCTAGCCCTACAAGCA	CTCTCAAAACTCATGAAACCCCTCCGTACC	763
QY	481	CATACTCGCTGGTGAGCCTAT	TTTAAATACACACCTCACTCGGCTCCATGAGTCTCAGCC	540
DB	764	CATACTCGCTGGTAGGCTAT	TTTAAATACACACCTCACTCGGCTCCATGAGTCTCGGCC	823
QY	541	CAAAAACCCCTACTAACTGT	TTGGATGTGCTCCCTCGCACTTCAGGCCCATCATTTCAATC	600
DB	824	CAAAAACCCCTACTAACTGT	TTGGATGTGCTCCCTCGCACTTCAGGCCCATCATTTCAATC	883
QY	601	CTGTGTCGAAACAATGGAA	CAACTTCAGCACAGAAATTAACACCACTCCGTTTATAGTA	660
DB	884	CTGTACCTGAAACAATGGAA	CAACTTCAGCACAGAAATTAACACCACTCCGTTTATAGTA	943
QY	661	GGACCTCTGTGTTTCCAACT	GTGAAATACCCATACCTCAAACTCAGCTGTGTAATAATTT	720
DB	944	GGACCTCTGTGTTTCCAACT	GTGAAATACCCATACCTCAAACTCAGCTGTGTAATAATTT	1003
QY	721	AGCAATACTATAGACAAACA	CCAGCTCCCAATGCATCAGTGGGTAAACACTCCACACGA	780
DB	1004	AGCAATACTATAGACAAACA	CCCACTCCCAATGCATCAGTGGGTAACTCTCCACACAA	1063
QY	781	ATAGTCTGCTACCTCAGGA	TATATTTTTGCTGTGTACTCTCAGCCTATCATTTGTTTG	840
DB	1064	ATAGTCTGCTACCTCAGGA	TATATTTTTGCTGTGTACTCTCAGCCTATCATTTGTTTG	1123
QY	841	AATGGCTCTTCAGAACTCT	ATGTGCTTCCTCTCAATTTAGTGCCCTCATGACCATCTAC	900
DB	1124	AATGGCTCTTCAGAACTCT	ATGTGCTTCCTCTCAATTTAGTGCCCTCATGACCATCTAC	1183
QY	901	ACTGAAACAAGATTTATACA	ATCATGTGTAACCTTAAGCCCAACAACAAAGAGTACCCATT	960
DB	1184	ACTGAAACAAGATTTATACA	ATCATGTGTAACCTTAAGCCCAACAACAAAGAGTACCCATT	1243
QY	961	CTTCTCTTTGTTATCAGAC	GGAGTGCTAGGCAGACTAGGTACTGGCAATTTGGCAGTATC	1020
DB	1244	CTTCTCTTTGTTATCAGAC	GGAGTGCTAGGTGCACTAGGTACTGGCAATTTGGCAGTATC	1303
QY	1021	ACAACCTCTACTCAGTTCT	ATACAAAATACTCTAAGAAATAATATGGTGAATGGAAACAG	1080
DB	1304	ACAACCTCTACTCAGTTCT	ATACAAAATACTCTAAGAAATAATATGGTGAATGGAAACAG	1363
QY	1081	GTCACTTGACTCCCTGGT	GCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140
DB	1364	GTCCGCACTCCCTGGT	GCACCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1423
QY	1141	CAAAATCGAAGACCTTTAGA	CTTGCTAACCCGCAAAAGAGGGGGAACCTGTTTATTTTA	1200
DB	1424	CAAAATCGAAGACCTTTAGA	CTTGCTAACCCGCAAAAGAGGGGGAACCTGTTTATTTTA	1483
QY	1201	GGAGAGAAACGCTGTTAT	TATGTTTAACTCAAGATTTGCATCGAAGAGTTAAAGAA	1260
DB	1484	GGGGAAGAAATGCTGTTAT	TATGTTTAACTCAAGATTTGCATCGAAGAGTTAAAGAA	1543
QY	1261	ATTTCGAGATCGAATACA	ATGTAGAGCAGAGGAGCTTTCAAAAACACCGAACGCTGGGGCTC	1320
DB	1544	ATTTCGAGATCGAATACA	ATGTAGAGCAGAGGAGCTTTCAAAAACACCGAACGCTGGGGCTC	1603
QY	1321	CTCAGCCAAATGGATG	CGCTGGGTTCTCCCTTCTTAGGACCTTAGCAGCTCTATATTTG	1380
DB	1604	CTCAGCCAAATGGATG	CGCTGGGTTCTCCCTTCTTAGGACCTTAGCAGCTCTATATTTG	1663
QY	1381	TTACTCTCTCTTGAC	CCCTGATCTTTTAACTCTTGTAACTGTTGCTCTCTCCAGAAAT	1440
DB	1664	CTACTCTCTCTTGAC	CCCTGATCTTTTAACTCTTGTAACTGTTGCTCTCTCCAGAAAT	1723
QY	1441	GAAGCTGTAAAGCTTAC	AGATGGTCTTCAAAATGGAAACCCCA	1481

QY	601	CCTGTTCTCTGAACAATGGAACAATCTTCAGACAGAAATAAACAACCACTTCGGTTTAACTA	660	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Db	884	CTGTGACCTGACAACTGGACAACTTCAGACAGAAATAAACAACCACTTCGGTTTAACTA	943	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
QY	661	GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCACCTGTGTAAATTT	720	1 (bases 1 to 2694)
Db	944	GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCACCTGTGTAAATTT	1003	Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G.,
QY	721	AGCAATACATAGACACACACAGCTCCCAATCGATCATCTAGTGGGTAAACACCTCCACACGA	780	Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.
Db	1004	AGCAATACATAGACACACACAGCTCCCAATCGATCATCTAGTGGGTAAACACCTCCACACAA	1063	The endogenous retroviral locus ERVWE1 is a bona fide gene involved
QY	781	ATAGTCTGCTACCTACCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCATCATCTGTTG	840	in hominoid placental physiology
Db	1064	ATAGTCTGCTACCTACCTCAGGAATATTTTTGTCTGTGGTACCTCAGCCATCATCTGTTG	1123	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
QY	841	AATGGCTCTTCAGAACTATGTCTCTCTCATCTTCTAGTCCCTATGACCATCTAC	900	14757826
Db	1124	AATGGCTCTTCAGAACTATGTCTCTCTCATCTTCTAGTCCCTATGACCATCTAC	1183	2 (bases 1 to 2694)
QY	901	ACTGAACAAGATTTATACAACTCATGTCTGACCTTAAGCCCAACAAAGAGTACCCATT	960	Mallet, F., Bouton, O. and Oriol, G.
Db	1184	ACTGAACAAGATTTATACAACTCATGTCTGACCTTAAGCCCAACAAAGAGTACCCATT	1243	Direct Submission
QY	961	CTTCTCTTTGTTATCAGACAGAGTGCTAGGACACTAGGTACTGGCAATGGCAGTATC	1020	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
Db	1244	CTTCTCTTTGTTATCAGACAGAGTGCTAGGACACTAGGTACTGGCAATGGCAGTATC	1303	CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
QY	1021	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTCGACATGGAACAG	1080	d'Italie, Lyon 69364 cedex 07, France
Db	1304	ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTCGACATGGAACAG	1363	Location/Qualifiers
QY	1081	GTCACTGACCTCCCTGGTCACTTTCGAAGATCAACTTAACTCCTACGACGAGTAGTCCTT	1140	1. .2694
Db	1364	GTCCCGACCTCCCTGGTCACTTTCGAAGATCAACTTAACTCCTACGACGAGTAGTCCTT	1423	/organism="Homo sapiens"
QY	1141	CAAAATCGAAGAGCTTTAGACTTCTAAACCGCAAAAGAGGGGACCTGTTTATTTTA	1200	/mol_type="genomic DNA"
Db	1424	CAAAATCGAAGAGCTTTAGACTTCTAAACCGCTGAAGAGGGGACCTGTTTATTTTA	1483	/isolate="individual 55 allele B"
QY	1201	GGAGAAACCGCTGTTTATGTTTAACTCAATCAGAAATGTCTACTGAGAAAGTTAAAGAA	1260	/db_xref="taxon:9606"
Db	1484	GGGAAAGAACTGTTTATGTTTAACTCAATCAGAAATGTCTACTGAGAAAGTTAAAGAA	1543	/chromosomes="7"
QY	1261	ATTGAGATCGAATACAAATGTAGACGAGAGGCTTCAAAACACCGAACGCTGGGCGCTC	1320	/map="7q21-q22"
Db	1544	ATTGAGATCGAATACAACTAGACGAGAGGCTTCAAAACACCTGGACCTCGGCGCTC	1603	/sex="female"
QY	1321	CTGAGCAATGGATGCGCTGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380	/cell_type="PBMC's"
Db	1604	CTGAGCAATGGATGCGCTGGTTCTCCCTTCTTAGGACCTCTAGCAGCTAATATTG	1663	/note="endogenous_virus: HERV-W"
QY	1381	TTACTCTCTTTGGACCTGTATCTTTAACTCCTCTGTTAAGTTGTCTTCCAGAAAT	1440	43. .57
Db	1664	CTACTCTCTTTGGACCTGTATCTTTAACTCCTCTGTTAAGTTGTCTTCCAGAAAT	1723	/note="splice acceptor site"
QY	1441	GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA	1481	284. .1900
Db	1724	GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGAGTCCAA	1764	/note="syncytin"
RESULT 24	AF52051952			/product="envelope glycoprotein"
LOCUS		2694 bp	DNA	/codon_start=1
DEFINITION		Homo sapiens individual 55 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.		/protein_id="AAQ17582.1"
ACCESSION	AF520520			/db_xref="GI:33411009"
VERSION	AF520520.1	GI:33411008		/translation="MALPYHFLFETLLPSFTLTAPPCCRCMTSSSPYQBEFLWRMORP
KEYWORDS				GNIDAPSTRSLKSGTPTFTATHMPCRYHSATLCMHANTHYMTGKMINPSCPGGLGV
SEGMENT				TVCWTYFTQVGDGGVQDQAREKHKEVISQLTRVHGTSSPYKGLDLXSLHETLRT
SOURCE				HTRVLSFNLTTLGLHEVSAQNTNCWICLPNFRPYVSIIPVQWNNFSTEINTSV
ORGANISM				LVLGVLNLEIHTSNLTQKFSNTTNTNSQCIRVTPPTQIIVCLPGGIFVCGTSA
				YGLNGSSECFSLFVPPMTIYEQDLVNYVSKPRKRVPIIPFVIGAGVLGALG
				TCIGITSTQFYVYKLSQELNGDMERVADSLVTLQDQNSLAAYVQLNRRDLTLTAE
				AGGCLFLGECSCYVQSGIVTEKVEIRDRIORRAEELRNTGPGULSOWMPWILP
				FLGPIAAIILLLLPFCPIFNLLVNFVSRIRIEAVKLQMEPKQSKTKIYRRPLDRPAS
				RSVDNDIKGTPFEISIAAFQPLLRPNAGSS"
				1705. .1720
				/note="splice acceptor site"
				1959. .2694
				ORIGIN
				Query Match 90.2%; Score 1335.4; DB 9; Length 2694;
				Best Local Similarity 93.9%; Pred. No. 0;
				Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
				QY 1 ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTACCCCTTTCCCTCTCACT 60
				Db 284 ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTTACCCTCTTCACTCTCACT 343
				QY 61 GCACCCCTCCCTGCTGTGTAACAACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
				Db 344 GCACCCCTCCATGCCGTGTATGACCATGCTCCCTTACCAAGAGTTTCTATGGAGA 403
				QY 121 ACGCGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAACTCC 180
				Db 404 ATGACGCGTCCCGAAATATTGATGCCCATCGTAGGAGTCTTTCTAAGGAAACCCCC 463
				QY 181 ACCTTCACGTCACACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
				Db 464 ACCTTCACGTCACACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 523
				QY 241 CATGCAAAATCACTATTATTGGACAGGAAATGATTAATCTTAGTTGTCTCTGGAGACTT 300
				Db 524 CATGCAAAATCACTATTATTGGACAGGAAATGATTAATCTTAGTTGTCTCTGGAGACTT 583

301 GGAGCCACTCTCTGTTGGACTTACTTACCCATACAGATATGTCTGATGGGGGTGAAT 360
 584 GGAGTCACTGTCTGTTGGACTTACTTACCCAACTGATATGTCTGATGGGGGTGAAT 643
 361 CAAGGTCAGCAAGAGAGAAAAAACAAGTAAAGAAAGCAATCTCCAACTGACCCGGGACAT 420
 644 CAAGATCAGGCAAGAGAGAAAAAACAAGTAAAGAAAGTAACTCTCCCAACTCACCCTGGTACAT 703
 421 AGACCCCTAGCCCTCAAGAGCACTAGTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 704 GGCACCTCTAGCCCTCAAGAGCACTAGTCTCTCAAACTACATGAAACCCCTCCGTACC 763
 481 CATACTCGCTGGTGGAGCTTAAATACCAACCTCTCACTGGGTCCATGAGGTCTCAGCC 540
 764 CATACTCGCTGGTGGAGCTTAAATACCAACCTCTCACTGGGTCCATGAGGTCTCAGCC 823
 541 CAAACCCCTACTAATCTGTTGGATGCTCTCCCTGCACTTCCAGGCAATACATTTCAATC 600
 824 CAAACCCCTACTAATCTGTTGGATGCTCTCCCTGCACTTCCAGGCAATATGTTTCAATC 883
 601 CTGTTCTCTGAACATGGAACACTTTCAGACAGAAATTAACACCACTTCGGTTTGTAGTA 660
 884 CTGTTCTCTGAACATGGAACACTTTCAGACAGAAATTAACACCACTTCGGTTTGTAGTA 943
 661 GGACCTCTGTTTCCAACTCTGGAATTAACCACTCAACCTCACTGTTGTAATAATTT 720
 944 GGACCTCTGTTTCCAACTCTGGAATTAACCACTCAACCTCACTGTTGTAATAATTT 1003
 721 AGCAATACTATAGACACACAGCTCCCAATGATGAGTGGTGAACACCTTCCACACGA 780
 1004 AGCAATACTATAGACACACAGCTCCCAATGATGAGTGGTGAACCTTCCACACAA 1063
 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGATCTCAGGCTATCATTTGTTG 840
 1064 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGATCTCAGGCTATCATTTGTTG 1123
 841 AATGGCTCTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 1124 AATGGCTCTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1183
 901 ACTGAAACAAGATTTATCAATCATGCTGATCACTAAGCCCAACAAACAAGAGTACCCATT 960
 1184 ACTGAAACAAGATTTATCAATCATGCTAATCTAAGCCCAACAAACAAGAGTACCCATT 1243
 961 CTTCCTTTTGTATCAGACGAGAGTGTCTAGGACAGATAGTGTCTGCTGCTGCTGCTGCT 1020
 1244 CTTCCTTTTGTATCAGACGAGAGTGTCTAGGACAGATAGTGTCTGCTGCTGCTGCTGCT 1303
 1021 ACAACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATAAATGTTGACATGGAACAG 1080
 1304 ACAACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATAAATGTTGGAACATGGAACGG 1363
 1081 GTCACTGACTCCCTGGTCACTTCAAGATCAACTTCACTCCCTAGCAGAGTATGCTCTT 1140
 1364 GTCCGCACTCCCTGGTCACTTCAAGATCAACTTCACTCCCTAGCAGAGTATGCTCTT 1423
 1141 CAAATCGAAGAGCTTTAGACTGTCTAACCACAAAGAGGGGAACTGTTTATTTT 1200
 1424 CAAATCGAAGAGCTTTAGACTGTCTAACCACAAAGAGGGGAACTGTTTATTTT 1483
 1201 GGAGAGAACCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1260
 1484 GGGGAAGAATGCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1543
 1261 ATTCAGATCGAATACAACTAGACGAGAGCTTCAAAACACCGCAACCTGCGGCGCTC 1320
 1544 ATTCAGATCGAATACAACTAGACGAGAGCTTCAAAACACCGCAACCTGCGGCGCTC 1603
 1321 CTCAGCCCAATGAGTCCCTGGTCTCTCCCTTTCTTAGGACCTCTAGCAGCTTATATTTG 1380
 1604 CTCAGCCCAATGAGTCCCTGGTCTCTCCCTTTCTTAGGACCTCTAGCAGCTTATATTTG 1663
 1381 TTACTCTCTTTGGACCCCTGTATCTTTAAACCTCTTGTAGTTTGTCTCTCTCCAGATTT 1440

1664 CTACTCCTCTTTGGACCTGTATCTTTAAACCTCTTTGTAACCTTTCTTCCAGATC 1723
 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 1724 GAAGCTGTAAAGCTACAAATGAGCCCAAGATGCACTCCAA 1764

RESULT 25
 AF520521S2 2694 bp DNA linear PRI 11-FEB-2004
 LOCUS Homo sapiens individual 6 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.
 DEFINITION AF520522
 ACCESSION AF520522
 VERSION AF520522.1 GI:33411012
 KEYWORDS
 SEGMENT
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2694)
 AUTHORS Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Wandrand, B.
 TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBMED 14757826
 REFERENCE 2 (bases 1 to 2694)
 AUTHORS Mallet, F., Bouton, O. and Oriol, G.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France

FEATURES
 source
 1. 2694
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="individual 6 allele A"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-q22"
 /sex="male"
 /cell_type="PBMC's"
 /note="Caucasian"
 endogenous_virus: HERV-W
 misc_feature 43..57
 /note="splice acceptor site"
 CDS 284..1900
 /note="syncytin"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="AAQ17583.1"
 /db_xref="GI:33411013"
 /translation="MALPYHIFLFTVLLPSFTLTAPPFCMTSSSPYQEPFLWRMORP
 GNIDAPSYRLSKGTPTFTATMPRNCYHSATLCMHANTHTYTKMNPSCPGGLGV
 TVCWYPTQMGSDGGVQDOAKHKVKEVISOLTRVHVGSIPVPEQMNFSIEINTSV
 HRLVSLPNTLTGLHEVSAQNPNCWICLPLNRPVYIPVPEQMNFSIEINTSV
 LTVGLVNSLFTHTSNLTCVKFSNTTNTTNSQCIKRWTPPTQIIVCLPSGIFVCGTSA
 YRCINGSSMCFSLFVLPMTIYEQDLNYSIKRPNKRPVLPFVIGAGVIGALG
 TGGTITTSQFYFKLSQELNGDMERVADSLVTLDQLNSLAAYVQLNRRALDLLTAE
 RGGTCLFLGECCEYVQSGIVTEKVEIKIRDIORAEELNTPGWGLLSQMWPILP
 FLGPLAIIILLGLFGPCIFNLVNFVSSRIEAVKLQMEPKMQSKTKIVRRPLDRPAS
 RSDVNDKIGTPPEISAAQFLLRNSAGSS"
 misc_feature 1705..1720
 /note="splice acceptor site"
 LTR 1959..2694

ORIGIN
 Query Match 90.2%; Score 1335.4; DB 9; Length 2694;
 Best Local Similarity 93.9%; Pred. No. 0;
 Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATATCTTTTCTCTTTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
Db 284 ATGGCCCTCCCTTATCATATCTTTTCTCTTTTACTGTTCTTACCCCTTTTCACTCTCACT 343
QY 61 GCACCCCTCCATGCTGCTGTGTAACAGTAGTCTCCCTTACAGAGTTTCTATGAAGA 120
Db 344 GCACCCCTCCATGCTGCTGTGTAACAGTAGTCTCCCTTACAGAGTTTCTATGGAGA 403
QY 121 AGCGGCTTCTGGAATATTCATGCCCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 404 ATGCAGCTGCCGGAATATTCATGCCCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC 463
QY 181 ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
Db 464 ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 523
QY 241 CATGAAATACCTATTATTTGGACAGGGAATAATGATTAATCTTAGTTGCTCTGGAGACTT 300
Db 524 CATGAAATACCTATTATTTGGACAGGGAATAATGATTAATCTTAGTTGCTCTGGAGACTT 583
QY 301 GGAGCCACTGCTGTGGACTTACTTACCCCATACAGTATCTCTGATGGGGGTGGAATT 360
Db 584 GGAGTCACTGTCTGTGGACTTACTTACCCCAACTGGTATGTCGTATGGGGGTGGAGTT 643
QY 361 CAAGGTCAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGGACAT 420
Db 644 CAAGATCAGCAAGAGAAAAACATGTAAAGAAAGTAATCTCCCACTCAACCCGGGTACAT 703
QY 421 AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAAACTACATGAACACCTCCGTACC 480
Db 704 GGCACTCTAGCCCTTACAAAGACTAGATCTCTCAAAACTACATGAACACCTCCGTACC 763
QY 481 CATACTCGCTGGTAGCTTATTAATACCACTTCACTCGCTCCATGAGTCTCAGCC 540
Db 764 CATACTCGCTGGTAGCTTATTAATACCACTTCACTCGCTCCATGAGTCTCGGCC 823
QY 541 CAAACCCCTACTAACTGTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
Db 824 CAAACCCCTACTAACTGTGTGGATGTGCTCCCTCGCACTTCAGGCCATATGTTCAATC 883
QY 601 CTGTGTCCTGAACAATGGAACAACCTTCAGACAGAAATAAACAACCTTCGTTTGTAGTA 660
Db 884 CTGTGTCCTGAACAATGGAACAACCTTCAGACAGAAATAAACAACCTTCGTTTGTAGTA 943
QY 661 GGACCTCTGTTTCCAACTGGAATAAACCCTACCTCAAACTCACTGCTGTAAATTT 720
Db 944 GGACCTCTGTTTCCAACTGGAATAAACCCTACCTCAAACTCACTGCTGTAAATTT 1003
QY 721 AGCAATACATAGACAAACCACTCCCAATGATCAGGTGGGTAAACCTCCCAACGA 780
Db 1004 AGCAATACATAGACAAACCACTCCCAATGATCAGGTGGGTAACTCTCCCAACAA 1063
QY 781 ATAGTCTGCTACCTCAGGAATATTTTTTGTGCTGTTGCTGACCTCAGCCTATCATTTG 840
Db 1064 ATAGTCTGCTACCTCAGGAATATTTTTTGTGCTGTTGCTGACCTCAGCCTATCGTTTGT 1123
QY 841 AATGGCTCTTCAGAACTATGCTTCTCTCATCTTAGTGCCCTCATGACCATCTAC 900
Db 1124 AATGGCTCTTCAGAACTATGCTTCTCTCATCTTAGTGCCCTCATGACCATCTAC 1183
QY 901 ACTGAACAAGATTATACAACTATGCTGATCTTAAGCCCAACAACAAGAGTACCCATT 960
Db 1184 ACTGAACAAGATTATACAACTATGCTGATCTTAAGCCCGCAACAAGAGTACCCATT 1243
QY 961 CTTCTCTTTTGTATCAGACAGAGTGTAGGACACTAGGTACTGGCATTTGGCATATC 1020
Db 1244 CTTCTCTTTTGTATAGGACAGAGTGTAGGACACTAGGTACTGGCATTTGGCGGTATC 1303
QY 1021 ACAACCTCTACTCAGTTCTTACAAACTATCTCAAGAAATAAATGGTACATGGAACAG 1080
Db 1304 ACAACCTCTACTCAGTTCTTACAAACTATCTCAAGAAATAAATGGGACATGGAACGG 1363
QY 1081 GTCACCTGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140

Db 1364 GTCGCCGACTCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1423
QY 1141 CAATATCAAGAGCTTTAGACTTGTACCCGCAAAAGAGGGGAACCTGTTTATTTT 1200
Db 1424 CAATATCAAGAGCTTTAGACTTGTACCCGCTGAAGAGGGGAACCTGTTTATTTT 1483
QY 1201 GGAGAAGAACCTGTTTATTTATTTAAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
Db 1484 GGGGAAGAACTGCTGTTTATTTATTTAAATCAATCCGGAATCGTCACCTGAGAAAGTTAAAGAA 1543
QY 1261 ATTGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCGAAGCTGGGGCCTC 1320
Db 1544 ATTGAGATCGAATACAACTAGACGAGAGAGCTTCGAAACACCTGGACCTGGGGCCTC 1603
QY 1321 CTCAGCCAAATCGATCGCTGGTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 1604 CTCAGCCAAATCGATCGCTGGATCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1663
QY 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTTCCAGAAAT 1440
Db 1664 CTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTAACTTGTCTCTTCCAGAAATC 1723
QY 1441 GAAGCTGTAAAGCTACAGATGCTCTTCAAAATGGAACCCCA 1481
Db 1724 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGAGTCCAA 1764

RESULT 26
AF520529S2
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SEGMENT
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AF520529S2
Homo sapiens individual 78 allele A, DNA linear
complete cds, and 3' long terminal repeat, complete sequence.
AF520530.1 GI:33411028
2 of 2
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2694)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
2 (bases 1 to 2694)
Mallet,F., Bouton,O. and Oriol,G.
Direct Submission
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 78 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="Caucasian
endogenous_virus: HERV-W"
43. .57
/note="splice acceptor site"
284. .1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17587.1"

FEATURES		d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers		1..2694
source		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/isolate="individual 71 allele A"
		/db_xref="taxon:9606"
		/chromosome="7"
		/map="7q21-q22"
		/sex="male"
		/cell_type="PBMC's"
		/note="African"
		endogenous_virus: HERV-W
misc_feature		43..57
		/note="splice acceptor site"
CDS		284..1900
		/note="syncytin"
		/codon_start=1
		/product="envelope glycoprotein"
		/protein_id="AAQ17603.1"
		/db_xref="GI:33411093"
		/translation="WALPYHIEFLTVLPSFTLTAPPRCMTSSSPYQEFILWNRQRP GNIDAPSVSLSKGPTFTAHMPNRCVHSATLQWANHNTYTKMINPSPCGGLGV TWQVYFTQGMDSGDGVQDOAREKHVKEVISOLTRVHGTSPPYKGLDLSKLHETLR HRLVSLFNTLTGLHEVSAQNPTNCWICLPIFRPYVISIPVEQNMFEINTISV LVGPLVSNLEIHTSNLTCVKFSNTYITNSOCIRWPTPTQIVCLPSGIFVCGTSA YRLNGSSMCPFLSELPVPMITYEQDLNYSVSKPRNKRVPILFPFVIGVLGALG TGIGTITTFQYFKLSQBLNGDMERVADSLVTLQDLNSLAUVLQNRRLDLTLAE RGCTFLGEECCYYVNOQGIIVTEKVKIEIRRIORAEELRNTGPWGLLSQMPWILP FLGPLAAIILLGLPGCIENLNVFVSRIEAVKLQMEPKMQSKTIVRRPLDRPASP RSDVDIKGTPPEEISAAQPLLRPNAGSS"
misc_feature		1705..1720
		/note="splice acceptor site"
LTR		1959..2694
ORIGIN		Query Match 90.2%; Score 1335.4; DB 9; Length 2694; Best Local Similarity 93.9%; Pred. No. 0; Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
Qy	1	ATGGCCCTCCCTTATCATATCTTTTCTTTTACTTGTTCTTCTTACCCCTTTTGGCTCTCACT 60
Db	284	ATGGCCCTCCCTTATCATATCTTTTCTTTTACTTGTTCTTTTACCCCTTTTCACTCTCACT 343
Qy	61	GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTACCAAGAGTTCTTATGAAGA 120
Db	344	GCACCCCTCCATGCTGCTGATGACAGTAGCTCCCTTACCAAGAGTTCTTATGAAGA 403
Qy	121	ACGGGCTTCTGGAATATTGATGCCCATATATAGAGAGTTTATCTAAGGGAACTCC 180
Db	404	ATGCAGCGTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACTCC 463
Qy	181	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCGCACTCTTTGATG 240
Db	464	ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCGCACTCTTTGATG 523
Qy	241	CATGCAAACTCATTTATGGACAGGAAATAGTTAATCTAGTTGTCCTGGAGGACTT 300
Db	524	CATGCAAACTCATTTATGGACAGGAAATAGTTAATCTAGTTGTCCTGGAGGACTT 583
Qy	301	GGAGCCACTGCTGTTGACCTTACTTCAACCATACCAAGTATGCTGATGGGGTGGAAAT 360
Db	584	GGAGTCACTGCTGTTGACCTTACTTCAACCATACCAAGTATGCTGATGGGGTGGAGTT 643
Qy	361	CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGGACAT 420
Db	644	CAAGATCAGGCAAGAGAAAAACATGTAAGAAAGTAATCTCCCAACTCACCCGGGTACAT 703
Qy	421	AGCACCTAGCCCTCAAGAGGACTGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db	704	GGCACCTTAGCCCTCAAGAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC 763
Qy	481	CATACTCGCTGGTGAGCCTATTTAATACCAACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db	764	CATACTCGCTGGTGAAGCCTATTTAATACCAACCCCTCACTCGGCTCCATGAGGTCTCGGCC 823
Qy	541	CAAAACCTACTACTGTTGGATGTCCTCCCTGACATTCAGGCCATACATTTCAATC 600
Db	824	CAAAACCTACTACTGTTGGATGTCCTCCCTGACATTCAGGCCATATGTTTCAATC 883
Qy	601	CCTGTTCTCTGAACAATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTA 660
Db	884	CCTGTACTGAAACATGGAACAACTTCAGCACAGAAATAAACACCACTTCGGTTTTAGTA 943
Qy	661	GGACCTCTTGTTCCTCAATCTGGAATAAACCATACCTCAAACTCACTGTTGTTAAATTT 720
Db	944	GGACCTCTTGTTCCTCAATCTGGAATAAACCATACCTCAAACTCACTGTTGTTAAATTT 1003
Qy	721	AGCAATACTATAGACACAACAGCTCCCAATGCAATCAGGTGGGTAAACCTCCACACAGA 780
Db	1004	AGCAATACTATAGACACAACCACTCCCAATGCAATCAGGTGGGTAACTCTCCACACAA 1063
Qy	781	ATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGGCTATCATTTGTTG 840
Db	1064	ATAGTCTGCCTACCTCAGGAATATTTTGTCTGTGTGCTACCTCAGGCTATCATTTGTTG 1123
Qy	841	AATGGCTCTTCAGAACTCATGTGCTTCTCTCATTTCTAGTGCCCTTATGACCATCTAC 900
Db	1124	AATGGCTCTTCAGAACTCATGTGCTTCTCTCATTTCTAGTGCCCTTATGACCATCTAC 1183
Qy	901	ACTGAACAAGATTTTATCAATCATGTCTACTAAGCCCAACAACAAAGAGTAGTACCCATT 960
Db	1184	ACTGAACAAGATTTTATCAATTAATGTCTATCTAAGCCCAACAACAAAGAGTAGTACCCATT 1243
Qy	961	CTTCCTTTGTTTATCAGACAGGAGTGTAGGACAGTATGATCTGGAATTTGGCAGTATC 1020
Db	1244	CTTCCTTTGTTTATAGGAGCAGGAGTGTAGTGTGCTAGTCTAGTCTGCGGTATC 1303
Qy	1021	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATATGTTGATGGAACAG 1080
Db	1304	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATATGTTGGAACAG 1363
Qy	1081	GTCACTGACTCCCTGCTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTACTCCTT 1140
Db	1364	GTCCGCGACTCCCTGCTCACTTCAAGATCAACTTAACTCCCTAGCAGCAGTACTCCTT 1423
Qy	1141	CAAAATCGAAGAGCTTTAGCTGTAACCGCAAAAGAGGGGAACTGTTTATTTTAA 1200
Db	1424	CAAAATCGAAGAGCTTTAGCTGTAACCGCTGAAGAGGGGAACTGTTTATTTTAA 1483
Qy	1201	GGAGAAAGCCTGTTTATGTTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA 1260
Db	1484	GGGGAAGAAATGCTGTTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1543
Qy	1261	ATTTCAGAGATCGAATAACAATGTAGACAGAGGAGCTTCAAAAACCGAACGCTGGGCGCTC 1320
Db	1544	ATTTCAGATCGAATAACAAGTAGACAGAGGAGCTTCGAAAACCTGGACCTGGGCGCTC 1603
Qy	1321	CTCAGCAATAGACCTCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db	1604	CTCAGCAATAGACCTCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTAATATTG 1663
Qy	1381	TTACTCCTCTTTGGACCTGTTATCTTTAACTCCTTGTAAAGTTGTCTCTTCCAGAAAT 1440
Db	1664	CTACTCCTCTTTGGACCTGTTATCTTTAACTCCTTGTAACTTGTCTCTTCCAGAAATC 1723
Qy	1441	GAACTGTAAAGCTACAGATGGTCTTTACAAATGGAAACCCCA 1481
Db	1724	GAACTGTAAAGCTACAAATGGAGCCCAAGATGCAAGTCCAA 1764
RESULT 28		
AX355872		
LOCUS		2930 bp DNA linear PAT 06-FEB-2002
DEFINITION		Sequence 1 from Patent WO0204678.
ACCESSION		AX355872

VERSION AX355872.1 GI:18620523
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Keith, J.C., McCoy, J.M. and M.I.S.
 TITLE Methods and compositions for diagnosing and treating preclampsia and gestational trophoblast disorders
 JOURNAL Patent: WO 0204678-A 1 17-JAN-2002;
 GENETICS INSTITUTE, INC. (US)
 FEATURES
 source 1..2930
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 CDS 930..2546
 /notes="unnamed protein product"
 /codon_start=1
 /protein_id="CAD23000.1"
 /db_xref="GI:18620524"
 /translation="MALPXYHIFLTVLLPSFTLTAPPKRCMTSSPYOEFLLMRMORP
 TNDVAFYSLSKGTPTFTAHMPNRCYHSALTCMHANTHYTGKMINPSCPGIGV
 GIVCWYFTQSMGSDGGVQDQAREKHVKEVISOLTRVHGTSYPYKGLIDLSKLHETLR
 HTLVSLENTLLGLHEVSAONPTNCWICLPLNFRPVSIPIVPEONNPNSTENITTSV
 LVGLVSNLEITHFTSHLTQVKNFTYNTNSQIRVPTPTOIVCLPSGIFVFCGSA
 YRLGNSSESSMFLSFLVPPMTIYTDQLYNYVISKPRNKRVPILFVVGAGVLGALG
 TGGGTTSTQFYKLSOELNGDMERVADSLVTLQDOLNSLAALVILNRRALLLTAE
 RGCTCLFBECCVYVQSGIVTEKVEIKRDRIOREAEELNTPGMLLSOMMPWTLIP
 FLGLAAIILLLLFGPCIFNLLVNFVSSRIEAVKLQMEPKMQSKTKIYRPLDRPASP
 RSDVNDIKGTTPPEISNAQPLLRPNAGSS"
 ORIGIN
 Query Match 90.2%; Score 1335.4; DB 6; Length 2930;
 Best Local Similarity 93.9%; Pred. No. 0;
 Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTATCATCTTTCTCTTACTGTTCTCTTACCCCTTTGGCTCTCACT 60
 DB 930 ATGGCCCTCCCTATCATCTTTCTCTTACTGTTCTCTTACCCCTTTGGCTCTCACT 989
 QY 61 GCACCCCTCCATGCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
 DB 990 GCACCCCTCCATGCTGTATGACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 1049
 QY 121 ACGGGCTTCTCGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
 DB 1050 ATGCAGGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAAACCCC 1109
 QY 181 ACCTTCACTGCCACACCATATGCCCGGCACTGCTATACCTGCGCACTTTGGCATG 240
 DB 1110 ACCTTCACTGCCACACCATATGCCCGGCACTGCTATACCTGCGCACTTTGGCATG 1169
 QY 241 CATCGAATCTCAATTGACAGGAAATGATTAATCTAGTGTCTGCGAGGACTT 300
 DB 1170 CATCGAATCTCAATTGACAGGAAATGATTAATCTAGTGTCTGCGAGGACTT 1229
 QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCATGATGTCTGATGGGGTGGAAAT 360
 DB 1230 GGAGTCACTGTCTGTGGACTTACTTCAACCATACCATGATGTCTGATGGGGTGGAGTT 1289
 QY 361 CAAGGTAGGCAAGAGAAAAACAAGTAAGGAAGCAATCTCCCAACTGACCCCGGGACAT 420
 DB 1290 CAAGATCAGGCAAGAGAAAAACATGTAAAGAAGTAATCTCCCAACTCACCCGGGTACAT 1349
 QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTCTCTCAAACTACATGAACCTCCGTACC 480
 DB 1350 GGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAACTACATGAACCTCCGTACC 1409
 QY 481 CATACTCGCCCTGGTGGAGCTATTTAATACCACTTCACTCGGCTCCATGAGGTCTCAGCC 540
 DB 1410 CATACTCGCCCTGGTAAAGCTATTTAATACCACTTCACTCGGCTCCATGAGGTCTCGGCC 1469

QY 541 CAAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCACTTCCAGGCCATACATTTCAATC 600
 DB 1470 CAAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCACTTCCAGGCCATATGTTTCAATC 1529
 QY 601 CTTGTTCTGTAACAATGGAACAACCTTACAGACAGAAATAAACAACCACTTCCGTTTGTAGTA 660
 DB 1530 CTTGTTCTGTAACAATGGAACAACCTTACAGACAGAAATAAACAACCACTTCCGTTTGTAGTA 1589
 QY 661 GGACCTCTGTTTCCAACTCTGAAATTAACCATATACCTCAAACTCACTGTGTAAATTT 720
 DB 1590 GGACCTCTGTTTCCAACTCTGAAATTAACCATATACCTCAAACTCACTGTGTAAATTT 1649
 QY 721 AGCAATATATAGACACAAACACAGCTCCCAATGCATCAGTGGGTAAACACTCCCAACGGA 780
 DB 1650 AGCAATATATAGACACAAACACAGCTCCCAATGCATCAGTGGGTAAACACTCCCAACGGA 1709
 QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGTACCTCAGCCTATCATTCGTTTG 840
 DB 1710 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGTACCTCAGCCTATCATTCGTTTG 1769
 QY 841 AATGGCTCTTCAAGATCTATGTGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
 DB 1770 AATGGCTCTTCAAGATCTATGTGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1829
 QY 901 ACTGAAACAAGATTATATACAAATCATGTCTGTAACCTTAAGCCCAACCAAAAGAGTACCCATT 960
 DB 1830 ACTGAAACAAGATTATATACAAATCATGTCTGTAACCTTAAGCCCAACCAAAAGAGTACCCATT 1889
 QY 961 CTTTCTTTTGTATCAGACGAGGAGTCTAGGACAGTACTAGTACTGGCAGTTCGCGATTC 1020
 DB 1890 CTTTCTTTTGTATCAGACGAGGAGTCTAGGAGTACTAGTACTGGCAGTTCGCGATTC 1949
 QY 1021 ACAACCTCTACTCAGTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 1080
 DB 1950 ACAACCTCTACTCAGTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 2009
 QY 1081 GTCACTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 DB 2010 GTCCGCGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059
 QY 1141 CAAAATCGAAGAGCTTTAGACTGCTTAAACCGCAAAAGAGGGGAGGAGGAGGAGGAGGAGG 1200
 DB 2070 CAAAATCGAAGAGCTTTAGACTGCTTAAACCGCTGAAAGAGGGGAGGAGGAGGAGGAGG 2129
 QY 1201 GGAGAAAGACGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1260
 DB 2130 GGAGAAAGACGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2189
 QY 1261 ATTTCGAGATCGAATAACAATGTAGACGAGGAGCTTCAAAACACCGAACCGCTGGGGCTC 1320
 DB 2190 ATTTCGAGATCGAATAACAATGTAGACGAGGAGCTTCAAAACACCGAACCGCTGGGGCTC 2249
 QY 1321 CTGAGCAATGAGATGCCCTGGGTTCTCCCTCTTTAGGAGCTCTAGCAGCTCTAATATTG 1380
 DB 2250 CTGAGCAATGAGATGCCCTGGGTTCTCCCTCTTTAGGAGCTCTAGCAGCTCTAATATTG 2309
 QY 1381 TTACTCTCTTTGAGCCCTGATCTTAACTCTCTTCTTAACTTAACTTAACTTAACTTAACTTAA 1440
 DB 2310 TTACTCTCTTTGAGCCCTGATCTTAACTCTCTTCTTAACTTAACTTAACTTAACTTAACTTAA 2369
 QY 1441 GAAGCTGTAAGGCTACAGATGGTCTTCAAAATGSAACCCCA 1481
 DB 2370 GAAGCTGTAAGGCTACAGATGGTCTTCAAAATGSAACCCCA 2410
 RESULT 29
 AF208161 2930 bp mRNA linear PRI 22-FEB-2000
 LOCUS Homo sapiens syncytin precursor, mRNA, complete cds.
 DEFINITION AF208161
 ACCESSION AF208161
 VERSION AF208161.1 GI:6760400
 KEYWORDS

SOURCE	Homo sapiens (human)	Db	1050	ATGCAGCGTCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTTAAAGGGAACCCCC	1109
ORGANISM	Homo sapiens	Qy	181	ACCTTCACTGCCACACCCCATATGCCCGCAATGCTATATAACTCTGCACTCTTTGTCATG	240
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Db	1110	ACCTTCACTGCCACACCCCATATGCCCGCAATGCTATATAACTCTGCACTCTTTGTCATG	1169
AUTHORS	1 (bases 1 to 2930) Mi.S., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L., LaVallie,E., Tang,X.Y., Edouard,P., Howes,S., Keith,J.C. Jr. and McCoy,J.M.	Qy	241	CATGCAATAACTCATTTATTTGGACAGGGAATAATGATTAACTCTAGTTTCTCTGGAGGACTT	300
TITLE	Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis	Db	1170	CATGCAATAACTCATTTATTTGGACAGGGAATAATGATTAACTCTAGTTTCTCTGGAGGACTT	1229
JOURNAL	Nature 403 (5771), 785-789 (2000)	Qy	301	GGAGCCACTGCTGTTGGACTTACTTACCACATACCATGATGCTCTGATGGGGTGGAAATT	360
MEDLINE	20155476	Db	1230	GGAGTCACTGCTGTTGGACTTACTTACCACCAACTGCTATGCTGATGGGGTGGAGTT	1289
PUBMED	10693809	Qy	361	CAAGTTCAGGCNAGAGAAACAAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGGACAT	420
REFERENCE	2 (bases 1 to 2930) Sha.M., Lee,X., Li,X., Veldman,G.M., Finnerty,H., Racie,L., LaVallie,E., Tang,X., Edouard,P., Howes,S., Keith,J.C. Jr. and McCoy,J.M.	Db	1290	CAAGATCAGGCAAGAGAAACAAATGTAAGAGATTAATCTCCCAACTCACCACCGGTACAT	1349
AUTHORS	Direct Submission Submitted (26-NOV-1999) Genetics Institute, 87 Cambridge Park Drive, Cambridge, MA 02140, USA Location/Qualifiers	Qy	421	AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAAAATACATGAAACCCCTCCGTACC	480
FEATURES	1..2930 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="AJ172-28" /tissue_type="testis" 1..2930 /tpt_family="Human endogenous retrovirus HERV-W" 1..929 930..2546 /note="envelope protein" /codon_start=1 /product="syncytin precursor" /protein_id="AAF28334.1" /gb_xref="GI:6760401" /translation="MALPVHIFLFTVLLPSFTLTAPPCCRCMTSSSPYOEFLLRMWRP GNIDAPYSLSKSGPTFTTAHTHMPNCHVSATLCHNANTHYWTGKMLNPSPCGGLGV TWCTVFTGTGSDGGVDQAREKHVKEVLSQLTRKVGTSIPYKGLDLSKHELRIT HYRLVSLFNTLTGLHEVAQNPCTWCICLPFRPYSIPVPEQWNPFSTEINTSV LVGPLVNEIHTSLNLTCKFSNTYTTNSQCIRWTPPTQIVCLPGSIFVFCGTS YRLNGSSSMFLSPLVPMITYTEQDLYNVIKPRNKRVPILPFVIGAVLGLGALG TGCGITLSTQFVKLSQELNGDMERVAADSLVTLQDLSLAVALVQNRRLADLLTAE RGCTCLFGECCYYNQSIGVTEKVEIKRDRIPRAEELRNTGPWGLLSQWMPILP FLGPLAAIILLDFGFCIFNLNVFVSSRIEAVKLQMEPMQSKIKIYRRPLDRPAS RSDVNDIKGTPPEISAAQPLLRPSAGSS" 930..989 990..2543 /product="syncytin" 1869..1880 /note="furin; cleavage site" /evidence=not_experimental 2043..2120 /note="Region: immunosuppressive region" 2235..2357 /note="transmembrane-region site" 2547..2930	Db	1350	GGCACCTTAGCCCCCTACAAAGGACTAGATCTCTCAAAAATACATGAAACCCCTCCGTACC	1409
repeat_region		Qy	481	CATACTCGCTGTGAGCCCTATTATTAACACCCCTCACTGGGCTCCATGAGGTCTCGGCC	540
5'UTR		Db	1410	CATACTCGCTGTGAGCCCTATTATTAACACCCCTCACTGGGCTCCATGAGGTCTCGGCC	1469
CDS		Qy	541	CAAAACCCCTACTAACTGTTGGATGCTCCCCCTGCACCTTCAGGCCATATGTTTCAATC	600
		Db	1470	CAAAACCCCTACTAACTGTTGGATGCTCCCCCTGCACCTTCAGGCCATATGTTTCAATC	1529
		Qy	601	CCTGTTCTCTGAAACAATGGAAACAATTCAGCACAGAAATAAACAACAATTCCTCGTTTAGTA	660
		Db	1530	CCTGTTCTCTGAAACAATGGAAACAATTCAGCACAGAAATAAACAACAATTCCTCGTTTAGTA	1589
		Qy	661	GGACCTCTGTTTCCAAATCTGGAATAAACCCATACCTCABACCTCACCTGTTGTAATAATT	720
		Db	1590	GGACCTCTGTTTCCAAATCTGGAATAAACCCATACCTCABACCTCACCTGTTGTAATAATT	1649
		Qy	721	AGCAATCTACTAGACACAAACAGCTCCCAATGTCATCAGGTGGTGAACACCTCCACACAGA	780
		Db	1650	AGCAATCTACTAGACACAAACAGCTCCCAATGTCATCAGGTGGTGAACACCTCCACACAGA	1709
		Qy	781	ATAGTCTCCCTACCCCTCAGGAATAATTTTGTCTGTGTGACTCTCAGCCCTATCATTTGTTG	840
		Db	1710	ATAGTCTCCCTACCCCTCAGGAATAATTTTGTCTGTGTGACTCTCAGCCCTATCATTTGTTG	1769
sig_peptide		Qy	841	AATGGCTCTTCAGAACTCTATGCTTCTCTCTCATTTCTTAGTGCCCTCTATGACCATCTAC	900
mat_peptide		Db	1770	AATGGCTCTTCAGAACTCTATGCTTCTCTCTCATTTCTTAGTGCCCTCTATGACCATCTAC	1829
misc_feature		Qy	901	ACTGAAACAAGATTTTATACAATCATGTCGTACCTAAGCCCCCAACAACAAGAGTACCCATT	960
misc_feature		Db	1830	ACTGAAACAAGATTTTATACAATCATGTCGTACCTAAGCCCCCAACAACAAGAGTACCCATT	1889
misc_feature		Qy	961	CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGCAGACTAGGTACTGGCACTGGCAGTATC	1020
3'UTR		Db	1890	CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGTGCACTAGGTACTGGCACTGGCAGTATC	1949
ORIGIN		Qy	1021	ACAACTCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAAACAG	1080
		Db	1950	ACAACTCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAAACAG	2009
		Qy	1081	GTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140
		Db	2010	GTCCGCGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	2069
		Qy	1141	CAAAATCGAAGACTTTAGACTTTGCACTTACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1200
		Db	2070	CAAAATCGAAGACTTTAGACTTTGCACTTACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2129
		Qy	1201	GGAGAAAGAACGCTGTTTATTATGTTTAACTCAATCAAGAAATGTTGTCACGTAGAAAGTAAAGAA	1260
		Db	2130	GGAGAAAGAACGCTGTTTATTATGTTTAACTCAATCCGGAATCGTCACTGAGAAAGTAAAGAA	2189

QY	1261	ATTGAGATCGAATAAATGTAGACGAGAGGCTTCAAAACACCGAAGCGTGGGGCTC	1320
DB	2190	ATTGAGATCGAATAAATGTAGACGAGAGGCTTCAAAACACCGAAGCGTGGGGCTC	2249
QY	1321	CTCAGCAATGATGATGCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	1380
DB	2250	CTCAGCAATGATGATGCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	2309
QY	1381	TTACTCTCTTTGGACCTGTATCTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGATT	1440
DB	2310	CTACTCTCTTTGGACCTGTATCTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGATT	2369
QY	1441	GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA	1481
DB	2370	GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA	2410
RESULT 30			
ARI77269			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1390; Conservative			
Score 90.2%;			
DB 6; Length 2946;			
Pred. No. 0;			
Indels 0; Gaps 0;			
Mismatches 91;			
1 ATGGCCCTCCCTTATCATATCTTTCTTTACTGTTCTTACCCCTTTACCCCTTTGGCTCTCACT			
928 ATGGCCCTCCCTTATCATATCTTTCTTTACTGTTCTTTTACCCCTTTTCACTCTCACT			
61 GCACCCCTCCATGCTGCTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA			
988 GCACCCCTCCATGCTGCTGTATGACCAAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA			
121 ACGGGCTTCTGGAATATGTATGCTCCCATCATATAGGAGTTTATCTAAGGGAATCTCC			
1048 ATGAGCGTCCCGGAATATGTATGCTCCCATCATATAGGAGTTTATCTAAGGGAATCTCC			
181 ACCTTCACTGCGCACACCATATATGCTCCCGCAACTGCTATAACTCTGCCACTTTTGGATG			
1108 ACCTTCACTGCGCACACCATATATGCTCCCGCAACTGCTATAACTCTGCCACTTTTGGATG			
241 CATGCAATACTCATATATGACAGGGAATATGATTAATCTAGTTGTCTGGAGGACTT			
1168 CATGCAATACTCATATATGACAGGGAATATGATTAATCTAGTTGTCTGGAGGACTT			
301 GGAGCCACTGTCTGTGGACTTACTTCCACCATACCATAGTATGCTGTAGTGGGGTGGAAAT			
1228 GGAGCCACTGTCTGTGGACTTACTTCCACCATACCATAGTATGCTGTAGTGGGGTGGAAAT			
361 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAATGACCGCGGAGAT			
1288 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAATGACCGCGGAGAT			
421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC			
DB	1348	GGCACTCTAGCCCTCAAAAGGACTAGTTCTCTCAAACTACATGAACCTCCGTACC	1407
QY	481	CATACTCGCTGGTGGAGCTTATTTAATACCACTCTACCTGGCTCCATGAGGTCTCAGCC	540
DB	1408	CATACTCGCTGGTGGAGCTTATTTAATACCACTCTACCTGGCTCCATGAGGTCTCAGCC	1467
QY	541	CAAAACCTACTAACTGTGTGATGCTCCCTCCCTGCACTTTAGGACCTCAATTTCAATC	600
DB	1468	CAAAACCTACTAACTGTGTGATGCTCCCTCCCTGCACTTTAGGACCTCAATTTCAATC	1527
QY	601	CCTGTTCTGAACTGGAACAACTTTCAGCAGCAAGAAATAAACACCACTTCCTGTTTAGTA	660
DB	1528	CCTGTTCTGAACTGGAACAACTTTCAGCAGCAAGAAATAAACACCACTTCCTGTTTAGTA	1587
QY	661	GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCCTGCTGTAAATTT	720
DB	1588	GGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCCTGCTGTAAATTT	1647
QY	721	AGCAATACTATAGACACCAACAGCTCCCAATGTCATCAGTGGGTAAACCTCCACACGA	780
DB	1648	AGCAATACTATAGACACCAACAGCTCCCAATGTCATCAGTGGGTAAACCTCCACACGA	1707
QY	781	ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCCTATCATTTGTTG	840
DB	1708	ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCCTATCATTTGTTG	1767
QY	841	ATAGTCTCTTCAAGATCTATGCTTCTCTCATCTTCTAGTCCCTTATGACCACTAC	900
DB	1768	ATAGTCTCTTCAAGATCTATGCTTCTCTCATCTTCTAGTCCCTTATGACCACTAC	1827
QY	901	ACTGAACAGATTTTATACAACTCATGTGTAACCTTAAGCCCAACAAAGAGTACCCATT	960
DB	1828	ACTGAACAGATTTTATACAACTCATGTGTAACCTTAAGCCCAACAAAGAGTACCCATT	1887
QY	961	CTTCTCTTTGTTATCAGCAGGAGTCTAGGACAGTAGTCTGCTGCTGCTGCTGCTGCTG	1020
DB	1888	CTTCTCTTTGTTATAGGACGAGGAGTCTAGGACAGTAGTCTGCTGCTGCTGCTGCTG	1947
QY	1021	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAAATTAATGGTGAACAG	1080
DB	1948	ACAACTCTACTCAGTTCTACTACAACTATCTCAAGAAATTAATGGTGAACAG	2007
QY	1081	GTCACTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
DB	2008	GTGCGGCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2067
QY	1141	CAAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCCAAGAGGGGGAACCTGTTATTTTA	1200
DB	2068	CAAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCCAAGAGGGGGAACCTGTTATTTTA	2127
QY	1201	GGAGAGAGAGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	1260
DB	2128	GGAGAGAGAGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	2187
QY	1261	ATTTCGAGATCGAATAACAATGTAGAGCAGAGAGCTTCAAAACACCGCAACCGTGGGGCTC	1320
DB	2188	ATTTCGAGATCGAATAACAATGTAGAGCAGAGAGCTTCAAAACACCGCAACCGTGGGGCTC	2247
QY	1321	CTCAGCAATGATGATGCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	1380
DB	2248	CTCAGCAATGATGATGCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG	2307
QY	1381	TTACTCTCTTTGGACCTGTATCTTTAAACCTCTCTGTTAAAGTTTGTCTCTTCCAGAAAT	1440
DB	2308	TTACTCTCTTTGGACCTGTATCTTTAAACCTCTCTGTTAAAGTTTGTCTCTTCCAGAAAT	2367
QY	1441	GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA	1481
DB	2368	GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA	2408

LOCUS BD232464 2946 bp DNA linear PAT 17-JUL-2003
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD232464
VERSION BD232464.1 GI:33042234
KEYWORDS JP 2002515234-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2946)
JACOBS,K., McCoy,J.M., Lavallie,E.R., Racie,L.A.C., Evans,C.,
Merberg,D., Mi,S. and Treacy,M.
Secreted proteins and polynucleotides encoding them
Patent: JP 2002515234-A 2 28-MAY-2002;
GENETICS INSTITUTE INC
OS Homo sapiens (human)
PN JP 2002515234-A/2
PD 28-MAY-2002
PF 17-MAY-1999 JP 2000549638
PR 18-MAY-1998 US 09/080478,20-OCT-1998 US 09/175928 PI
KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A COLLINS PI
RACIE,
PI CHERYL EVANS,DAVID MERBERG,SHA MI,MAURICE TREACY PC
C12N15/09,A61K38/00,A61K45/00,A61K48/00,A61P7/00,A61P17/02, PC
A61P29/00,
PC A61P35/00,A61P37/02,A61P43/00,C07K14/47,C12N5/10,C12P21/02, PC
C12Q1/02,
PC C12Q1/68,C12N15/00,A61K37/02,C12N5/00
CC Secreted proteins and polynucleotides encoding them FH Key
Location/Qualifiers
FT source 1..2946
FT Location/Qualifiers
1..2946
/organism='Homo sapiens (human)'.
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 90.2%; Score 1335.4; DB 6; Length 2946;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTATCATCTTTCTTTCTTACTGTTCTTACCCGCTTTGCTCTCACT 60
DB |||||||
QY 928 ATGGCCCTCCCTATCATATTTTCTTTCTTACTGTTCTTTTACCCCTTTTCACTCTCACT 987
DB |||||||
QY 61 GCACCCCTCCATGCTGTGTACACACGAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
DB |||||||
QY 988 GCACCCCTCCATGCGCTGTATGACGAGTAGCTCCCTTACCAAGAGTTTCTATGGAGA 1047
DB |||||||
QY 121 ACGGCGGTCCTGGAAATATTGATGCGCCNCTATATAGGAGTTTATCTAAGGGAACCTCC 180
DB |||||||
QY 1048 ATGCAGCGTCCCGGAAATATTGATGCGCCATCGTATAGAGTCTTTCTAAGGGAACCCCC 1107
DB |||||||
QY 181 ACCTTCACCTGCCACACCCATGCGCCGCACTGTATACTCTCCACTCTTTTGCATG 240
DB |||||||
QY 1108 ACCTTCACCTGCCACACCCATGCGCCGCACTGTATCACTCTGCCACTCTTTCATG 1167
DB |||||||
QY 241 CATGCCAAATCACTCAITATTGGACAGGGAATAATCTTAACTTGTCTCTGAGGACTTT 300
DB |||||||
QY 1168 CATGCCAAATCACTTATTGGACAGGGAATAATCTTAACTTGTCTCTGAGGACTTT 1227
DB |||||||
QY 301 GGAGCCACTGCTGTTGGACTTACTTACCACCATACAGTATGTCATGCGGGGTGAATTT 360
DB |||||||
QY 1228 GGAGTCACTGCTGTTGGACTTACTTACCACCAACTGGTATGTCATGCGGGGTGGAGTT 1287
DB |||||||
QY 361 CAAGGTTCAGCAAGAGAAAAAAGTAAAGGAAGCAATCTCCAACTGACCCCGGGGACAT 420
DB |||||||
QY 1288 CAAGATCAGGCAAGAGAAAAAATGTAAGAAGAGTAACTCTCCAACTCACCCTGGGTACAT 1347
DB |||||||
QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
DB |||||||

1348 GGCACCTCTAGCCCTCAAAAGGAGTACTCTCTCAAAACTTACATGAACCCCTCCGTACC 1407
QY |||||||
481 CATACTCGCTGGTGGAGCCTATTTAATAGCACCTCTCACTCGGCTCCATGAGGTCTCAGCC 540
DB |||||||
1408 CATACTCGCTGGTGGAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCGGCC 1467
QY |||||||
541 CAAACCCCTACTAATCTGTTGGATGTGCTCCCTCGCATCTTCAGGCCATACATTTCAATC 600
DB |||||||
1468 CAAACCCCTACTAATCTGTTGGATATGCTCCCTCGCACTTCAGGCCATATGTTTCAATC 1527
QY |||||||
601 CCTGTTCTGTAACAAATGGAACTCTGAGCAGAGAAATAACACACTTCCTCGTTTGTAGTA 660
DB |||||||
1528 CCTGTAGCTGAACAAATGGAACTTCAGCAGAGAAATAACACACTTCCTCGTTTGTAGTA 1587
QY |||||||
661 GGACCTCTGTTTCCCAATCTGGAATAAACCATACCTCAAACTCCTGCTGTGTAATAATTT 720
DB |||||||
1588 GGACCTCTGTTTCCCAATCTGGAATAAACCATACCTCAAACTCCTGCTGTGTAATAATTT 1647
QY |||||||
721 AGCAATACTATAGACACAAACAGCTCCCAATGCAATGCAATGAGGTGGGTAAACACTCCACAGGA 780
DB |||||||
1648 AGCAATACTATACACAAACCACTCCCAATGCAATGCAATGAGGTGGGTAACTCTCCACACAA 1707
QY |||||||
781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB |||||||
1708 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1767
QY |||||||
841 AATGCTCTTCCAGAACTCTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB |||||||
1768 AATGCTCTTCCAGAACTCTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1827
QY |||||||
901 ACTGAACAAGATTTATACAATCATGTGTGCTAAGCCCCACACAAAGAGTAGTCCCAATT 960
DB |||||||
1828 ACTGAACAAGATTTATACAATCATGTCTAAGCCCCACACAAAGAGTAGTCCCAATT 1887
QY |||||||
961 CTTCTCTTTGTTATCAGAGCAGGAGTCTAGGAGACTAGTACTGCGCATGCGCATGGCAGTAGTC 1020
DB |||||||
1888 CTTCTCTTTGTTATAGAGCAGGAGTGTCTAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGTG 1947
QY |||||||
1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGCTGACATGGAAACAG 1080
DB |||||||
1948 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGCGGACATGGAAACGG 2007
QY |||||||
1081 GTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCTCTAGCAGCAGTAGTCTCTT 1140
DB |||||||
2008 GTGCGGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCTCTAGCAGCAGTAGTCTCTT 2067
QY |||||||
1141 CAAATCGAAGAGCTTTAGACTTGTAAACCGCAAGAGGGGGAACCTGTTTATTTTATA 1200
DB |||||||
2068 CAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTATA 2127
QY |||||||
1201 GGAGAAGAACGCTGTTTATTTATTTAATCAATCCAGAATTTGCTACTGAGAAAGTTAAAGAA 1260
DB |||||||
2128 GGGGAAGAAATGCTGTTTATTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 2187
QY |||||||
1261 ATTCAGATTCGAATCAATGTAGACAGAGGCTTCAAACACCGAAGCTGGGCGCTC 1320
DB |||||||
2188 ATTCAGATTCGAATCAATCGTAGACAGAGGCTTCGAAACACCTGGACCTTGGGCGCTC 2247
QY |||||||
1321 CTCAGCCTAGAGTCCCTGGGTTCTCCCTTTCTTTAGGACCTCTAGCAGCTCTAATATTATG 1380
DB |||||||
2248 CTCAGCCTAGAGTCCCTGGGTTCTCCCTTTCTTTAGGACCTCTAGCAGCTAATATTATG 2307
QY |||||||
1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTTGTGTTAGTTGTCTCTTCCAGAAATT 1440
DB |||||||
2308 CTACTCTCTTTGGACCTGTATCTTTAACTCTCTTGTGTTAGTTGTCTCTTCCAGAAATC 2367
QY |||||||
1441 GAAGCTGTAAAGCTACAGATGCTCTTACAATGGAAACCCCA 1481
DB |||||||
2368 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAATGCCAA 2408
QY |||||||

LOCUS BD086133 2946 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD086133
VERSION BD086133.1 GI:22631743
KEYWORDS JP 2001524490-A/2.
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2946)
AUTHORS Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A.C., Evans,C.,
Merberg,D. and Treacy,M.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patent: JP 2001524490-A 2 04-DEC-2001;
COMMENT GENETICS INSTITUTE INC
OS Unidentified
FN JP 2001524490-A/2
PD 04-DEC-2001
PF 17-NOV-1998 JP 2000522128
PR 21-NOV-1997 US 08/976110,18-MAY-1998 US 09/080478 PR
20-OCT-1998 US 09/175928
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A COLLINS
PI RACIE,
PI CHERYL EVANS,DAVID MERBERG,MAURICE TREACY
PC C07K14/47,C12N5/10,C12N15/09,C12P21/02,C12N5/00,C12N15/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Secreted proteins and polynucleotides encoding them FH Key
FT source Location/Qualifiers
FT 1..2946 /organism='Unidentified'.
1..2946 Location/Qualifiers
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

Query Match 90.28; Score 1335.4; DB 6; Length 2946;
Best Local Similarity 93.94; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATATCTTTCTCTTACTTGTCTCTTACCCCTTTTCGCTCTCACT 60
DB 928 ATGGCCCTCCCTTATCATATCTTTCTCTTACTTGTCTCTTACCCCTTTTCGCTCTCACT 987
QY 61 GCACCCCTCCATGCTCTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 120
DB 988 GCACCCCTCCATGCTCTGTATGACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAGA 1047
QY 121 ACGGGCTCTCGAATATTTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAATCTCC 180
DB 1048 ATGCAGCGTCCCGAATATTTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAATCTCC 1107
QY 181 ACCTTCACTGCCACACCATATATGCCCGCAACTGCTATAACTCTGCCACTCTTTTGCATG 240
DB 1108 ACCTTCACTGCCACACCATATATGCCCGCAACTGCTATAACTCTGCCACTCTTTTGCATG 1167
QY 241 CATGCAAACTACTATTTATGACAGGGAATAATGATTAATCTAGTTGCTCTGAGGACTT 300
DB 1168 CATGCAAACTACTATTTATGACAGGGAATAATGATTAATCTAGTTGCTCTGAGGACTT 1227
QY 301 GGAGCCACTGCTGTGGACTTACTTACCCTATACCATGATGCTGATGGGGTGAATT 360
DB 1228 GGAGTCACTGCTGTGGACTTACTTACCCTATACCATGATGCTGATGGGGTGGAGTT 1287
QY 361 CAAGTCAGGCAAGAGAAAACAAGTAAGGAAGCAATCTCCCAACTGACCCCGGGGACAT 420
DB 1288 CAAGATCAGGCAAGAGAAAACAATGTAAGAGAGTATCTCCCAACTCACCCTGGGTACAT 1347
QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
DB 1348 GGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 1407

QY 481 CATACTCGCTGGTGAGCCTATTTTAATACCAACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB 1408 CATACTCGCTGGTGAAGCCTATTTTAATACCAACCCCTCACTGGGCTCCATGAGTCTCGGCC 1467
QY 541 GAAACCCCTACTAACTGTTGGATGTCCTCCCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
DB 1468 GAAACCCCTACTAACTGTTGGATGTCCTCCCTCCCTGAACTTCAGGCGCATATGTTTCAATC 1527
QY 601 CTGTTTCTGAAACAATGGAACAATCTTCAGCAGAGAAATAAACAACCACTTCCTCGTTTGTGTA 660
DB 1528 CTGTTACCTGACATGGAACAATCTTCAGCAGAGAAATAAACAACCACTTCCTCGTTTGTGTA 1587
QY 661 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCACTGCTGTGTAATTTT 720
DB 1588 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCACTGCTGTGTAATTTT 1647
QY 721 AGCAATACCTATAGACACACACACAGCTCCCAATCATCAGGTGGGTAAACACTCCCAACACA 780
DB 1648 AGCAATACCTATAGACAA 1707
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGTTACCTCAGCCTATATCTGTTTGT 840
DB 1708 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGTTACCTCAGCCTATATCTGTTTGT 1767
QY 841 AATGGCTCTTCAGAACTCTATGCTCTCTCTATCTTCTTGTAGTGGCCCTATGACCATCTAC 900
DB 1768 AATGGCTCTTCAGAACTCTATGCTCTCTCTATCTTCTTGTAGTGGCCCTATGACCATCTAC 1827
QY 901 ACTGAACAAGATTATACAACTCATGCTGCTTAAGCCCCCACAACAAAGAGTACCCATT 960
DB 1828 ACTGAACAAGATTATACAACTCATGCTGCTTAAGCCCCCACAACAAAGAGTACCCATT 1887
QY 961 CTTCTCTTTTGTATCAGACGAGGTGCTAGGAGACTAGGTACTGGCACTGGCACTGGCACTATC 1020
DB 1888 CTTCTCTTTTGTATCAGACGAGGTGCTAGGAGACTAGGTACTGGCACTGGCACTGGCACTATC 1947
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGAATGGAACAG 1080
DB 1948 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGAATGGAACAG 2007
QY 1081 GTCACTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 2008 GTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2067
QY 1141 CAAATCGAAGAGCTTTAGACTTTGCTAACCGCAAGAGAGGGGAGACCTGTTTATTTTAA 1200
DB 2068 CAAATCGAAGAGCTTTAGACTTTGCTAACCGCAAGAGAGGGGAGACCTGTTTATTTTAA 2127
QY 1201 GGAAGAGAGAGCTGTTTATGTTTAAATCAATCAGAAATGCTCACTGAGAAAGTTAAGAA 1260
DB 2128 GGGGAAGAGAGCTGTTTATGTTTAAATCAATCAGAAATGCTCACTGAGAAAGTTAAGAA 2187
QY 1261 ATTCGAGATCGAATCAATCTAGACGAGAGAGCTTCAAAACACCGCAACCTCGGGGCTC 1320
DB 2188 ATTCGAGATCGAATCAATCTAGACGAGAGAGAGCTTCAAAACACCGCAACCTCGGGGCTC 2247
QY 1321 CTCAGCAATGAGTGGCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAATATG 1380
DB 2248 CTCAGCAATGAGTGGCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAATATG 2307
QY 1381 TTAATCTCTCTTTGGAGCCCTGATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGATT 1440
DB 2308 CTACTCTCTTTGGAGCCCTGATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGATT 2367
QY 1441 GAAGCTGTAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481
DB 2368 GAAGCTGTAAGCTACAGATGGAGCCCAAGATGAGTCCAA 2408

RESULT 33
AF506835
LOCUS
DEFINITION Homo sapiens enverin mRNA, partial cds.

AF506835 1624 bp mRNA linear
PRI 20-MAY-2002

[illegible]

Db 1405 CTACTCTCTTTGGACCTGTATCTTTAAACCTCTCTTTTAACTTTGTCTTCTCCAGATC 1464
 QY 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAATGGAACCCCA 1481
 Db 1465 GAAGCTGTAAAGCTACAGATGGGACCCCAAGATGAGTCCAA 1505

RESULT 34
 AF513360 1860 bp mRNA linear PRI 05-JUN-2002
 LOCUS Homo sapiens enverin mRNA, complete cds.
 DEFINITION AF513360
 ACCESSION AF513360.1 GI:21326140
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1860)
 AUTHORS Alliel,P.M., Perin,J.P., Pierig,R., Nussbaum,J.L., Menard,A. and Rieger,F.
 TITLE Endogenous retroviruses and multiple sclerosis. Part 2: HERV-7q and its env transcripts
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 312, 857-863 (1998)
 REFERENCE 2 (bases 1 to 1860)
 AUTHORS Alliel,P.M., Perin,J.P., Goudou,D., Bitoun,M., Robert,B. and Rieger,F.
 TITLE The HERV-W/7q family in the human genome. Potential for protein expression and gene regulation
 JOURNAL Cell Mol. Biol. 48 (2), 213-217 (2002)
 MEDLINE 21985840
 PUBMED 1190458
 REFERENCE 3 (bases 1 to 1860)
 AUTHORS Alliel,P.M., Goudou,D., Perin,J.P. and Rieger,F.
 TITLE Direct Submission
 JOURNAL Submitted (18-MAY-2002) U-488, INSERM, 80, rue du General Leclerc, Le Kremlin-Bicetre 94270, France

FEATURES
 source
 1..1860
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21"
 /tissue_type="placenta"
 59..1675
 /codon_start=1
 /product="enverin"
 /protein_id="AAM47599.1"
 /db_xref="GI:21326141"
 /translation="MALPYHIFLFTVLLPSFTLTAPPPRCMTSSSPYQFLWRMQRP
 GNDAPSYRLSKGTPTFTAHMPNRYHSATLCMHANTHYWTGRKMINSPCGGLV
 TVCYTYFTQMGSDGGVQDQAREKHVKEVISQLTRVHGTSPPYKGLDLKSLHETLRT
 HTRLVSLFNTTTLGLHVSQNPNCWICLPNRPVXSIPVPEQMNFNTEINTTSV
 LVGPLVNLRTHTSNLTCKVFSNTVTNCSQIRWTPPTQIVCLPSGIPFVCGTSA
 YRLGNSSEKCFSLFVPPMTIYEQDLNLYVLSKPNKRVPLLPVIGAGVLGALG
 TRLGGLTTSFYKLSQBLNMDERVADSLVTLQDLNLSAAVLQNRALDLITRE
 RGCTFLGEECCYYNQSGIVTEKVEIRDRIOQRRAEELRNTGFWLLSQWMPWPLP
 FLGPLAAIILLFLPGPCIFNLLVNFVSRIEAVKLOMEPKMQSKTKIYRRLDRPAS
 RSDVNDIKGTPPEISAAQPLLRNSAGSS"

ORIGIN
 Query Match 90.1%; Score 1333.8; DB 9; Length 1860;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATCATCTTCTCTTACTGTCTTACCCCTTCCGCTCTCACT 60
 Db 59 ATGGCCCTCCCTTATCATATTTTCTTCTTACTGTCTTACCCCTTCTTCACT 118
 QY 61 GCACCCCTCCATGCTGCTATACCAACCAAGTAGCTCCCTTACCAAGAGTTTCTATGA 120
 Db 119 GCACCCCTCCATGCGCGTGTATGACCAGTAGCTCCCTTACCAAGAGTTTCTATGA 178

121 ACGCGCTTCTTGGAAATATTGATGCCCATATAGGAGTTTATCTAAGGAAACTCC 180
 179 ATGACGCTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGAAACCC 238
 181 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATAACTCTGCCACTCTTTGCATG 240
 239 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATCACTCTGCCACTCTTTGCATG 298
 241 CATGCAAAATACATTTATTGGACAGGAAATGATTAACTCTAGTTGTCTCGAGGACTT 300
 299 CATGCAAAATACATTTATTGGACAGGAAATGATTAACTCTAGTTGTCTCGAGGACTT 358
 301 GGAGCCACTGTCTGTGGACTTTACTTCAACCATACCAGTATGTCTGATGGGGTGGAAAT 360
 359 GGAGTCACTGTCTGTGGACTTTACTTCAACCATACCAGTATGTCTGATGGGGTGGAGTT 418
 361 CAAGGTGAGCAAGAGAAACAAAGTAAGAGAGCAATCTCCCACTGAGCCCGGGACAT 420
 419 CAAGATCAGGCAAGAGAAACAAAGTAAGAGAGTAACTCTCCCACTGAGCCCGGGTACAT 478
 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCTCCGATACC 480
 479 GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAACCTCACTGGGCTCCATGAGGTCTCGGCC 538
 481 CATACTCGCTGTGAGCCCTATTAAATACACCCCTCACTGGGCTCCATGAGGTCTCAGCC 540
 539 CATACTCGCTGTGAGCCCTATTAAATACACCCCTCACTGGGCTCCATGAGGTCTCGGCC 598
 541 CAAAAACCTACTAACTGTGGATGTCCTCCCTCGACTTTCAGGCGCATACATTTCAATC 600
 599 CAAAAACCTACTAACTGTGGATATGCTCCCTCGAACTTTCAGGCGCATATGTTTCAATC 658
 601 CCTGTTCTTGAACAATGGAACAACTTTCAGCAGAGAAATAAACACCCTCCGTTTTAGTA 660
 659 CCTGTACCTTGAACAATGGAACAACTTTCAGCAGAGAAATAAACACCCTCCGTTTTAGTA 718
 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCCCTGTGTAAATTT 720
 719 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCCCTGTGTAAATTT 778
 721 AGCAATATATAGACACCAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
 779 AGCAATATATAGACACCAACAGCTCCCAATGATCAGGTGGGTAACTCTCCACACAA 838
 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTTG 840
 839 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTTG 898
 841 AATGGCTCTTCAAGATCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 899 AATGGCTCTTCAAGATCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 958
 901 ACTGAACAAGATTTATACATATGCTGTAAGCTTAAAGCCCAACAAAGAGTACCCATT 960
 959 ACTGAACAAGATTTATACATATGCTGTAAGCTTAAAGCCCAACAAAGAGTACCCATT 1018
 961 CTTCTCTTTGTTTATCAGACGAGGTGCTAGGACAGTACTAGTCTGCGATTGGCAGTATC 1020
 1019 CTTCTCTTTGTTTATGAGGACGAGGTGCTAGGTCACATAGTCTGCGATTGGCAGTATC 1078
 1021 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGCTGATGGAACAG 1080
 1079 ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAAATGCTGATGGAACAG 1138
 1081 GTCACTGACTCCCTGGTCACTTGCAGAGTCAACTTAACTCCCTAGCAGAGTATGCTCT 1140
 1139 GTGCCGACTCCCTGGTCACTTGCAGAGTCAACTTAACTCCCTAGCAGAGTATGCTCT 1198
 1141 CAAAATCGAGAGCTTTAGACTTTGCTTAACCGCAAGAGGGGAACTGTTTATTTTAA 1200
 1199 CAAAATCGAGAGCTTTAGACTTTGCTTAACCGCTGAAGAGGGGAACTGTTTATTTTAA 1258

QY	1201	GGAGAAGAACGCTGTTATTTATGTTAAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA	1260
Db	1259	GGGAAGAAGTACGCTGTTACTATGTTAAATCAATCCGAATCGTCACCTGAGAAAGTTAAAGAA	1318
QY	1261	ATTCCAGATCGAATCAATGTTAGACAGAGGAGCTTCAAAACACCGAAGCTGGGGCCTC	1320
Db	1319	ATTCCAGATCGAATCAACGTTAGACAGAGGAGCTTGGAAACATCGGACCTGGGGCCTC	1378
QY	1321	CTCAGCCAAATGATGCCCTGGGTTCTCCCTTTCTTAGGACCTCTTAGCAGCTCTAATATG	1380
Db	1379	CTCAGCCAAATGATGCCCTGGGTTCTCCCTTTCTTAGGACCTCTTAGCAGCTCTAATATG	1438
QY	1381	TTACTCCTCTTTGGACCTGATCTTTTAACCTCCTTTGTTAAGTTTGTCTCTCCAGAAAT	1440
Db	1439	CTACTCCTCTTTGGACCTGATCTTTTAACCTCCTTTGTTAAGTTTGTCTCTCCAGAAAT	1498
QY	1441	GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA	1481
Db	1499	GAAGCTGTAAAGCTACAAATGAGCCCAAGATGCAATCCAA	1539
RESULT 35			
BD221827		2055 bp DNA linear PAT 17-JUL-2003	
LOCUS		Nucleic sequence and deduced protein sequence family with human	
DEFINITION		endogenous retroviral motifs, and their uses.	
ACCESSION	BD221827		
VERSION	BD221827.1	GI:33031597	
KEYWORDS	JP 2002518051-A/22.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Alliel,P.M., Perin,J.P. and Rieger,F.		
TITLE	Nucleic sequence and deduced protein sequence family with human		
JOURNAL	endogenous retroviral motifs, and their uses		
COMMENT	Patent: JP 2002518051-A 22 JUN-2002; INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE OS Homo sapiens (human) PN JP 2002518051-A/22 PD 25-JUN-2002 PF 23-JUN-1999 JP 2000556036 PR 23-JUN-1998 FR 98/07920 PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC C12N15/09, A01K67/027, A61K31/711, A61K39/21, A61K48/00, A61P21/00, PC A61P25/00, PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC Nucleic sequence and deduced protein sequence family with CC human endogenous CC retroviral motifs, and their uses FH Key Location/Qualifiers FT CDS Location/Qualifiers		
FEATURES			
source	1..2055	/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match	90.1%;	Score 1333.8;	DB 6; Length 2055;
Best Local Similarity	93.8%;	Pred. No. 0;	
Matches 1389;	Conservative 0;	Mismatches 92;	Indels 0; Gaps 0;
QY	1	ATGGCCCTCCCTTATCATATCTTTTCTCTTTACTGTTCTCTTACCCCTTTTCGCTCTCACT	60
Db	391	ATGGCCCTCCCTTATCATATCTTTTCTTTACTGTTCTTTTACCTCTTTCACTCTCACT	450
QY	61	GCACCCCTCCATGCTGTACAAACAGTAGCTCCCTTACCAGAGTTTCTATGAGA	120
Db	451	GCACCCCTCCATGCTGTATGACCAAGTAGCTCCCTTACCAGAGTTTCTATGAGA	510
QY	121	ACGGGGCTTCTCGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAAACTCC	180

Db	511	ATGAGCGTCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACCCCC	570
QY	181	ACCTTCACTGCCACACCCATATGCCCGCAACGTCTTAACCTGCACTCTTTTGATG	240
Db	571	ACCTTCACTGCCACACCCATATGCCCGCAACGTCTTAACCTGCACTCTTTTGATG	630
QY	241	CATGCAATCTCAATATTGGACAGGGAATGATTAACTCTAGTTGTCCTGGAGGACTT	300
Db	631	CATGCAATCTCAATATTGGACAGGGAATGATTAACTCTAGTTGTCCTGGAGGACTT	690
QY	301	GGAGCCACTGTCTTGGACTTACTTCAACCATACAGTATGTCTGATGGGGTGGAAAT	360
Db	691	GGAGTCACTGTCTTGGACTTACTTCAACCATACAGTATGTCTGATGGGGTGGAGTT	750
QY	361	CAAGGTGAGCAAGAGAAAACAAGTAAGGAGCAATCTCCCAACTGACCCCGGGACAT	420
Db	751	CAAGATCAGGCAAGAGAAAACAAGTAAGGAGCAATCTCCCAACTGACCCCGGGTACAT	810
QY	421	AGCACCCCTAGCCCTACAAAGAGTACTTCTCTCAAACTACATGAAACCTCCGTACC	480
Db	811	GGCACTCTAGCCCTACAAAGAGTACTTCTCTCAAACTACATGAAACCTCCGTACC	870
QY	481	CATACCTCGCTGGTGAAGCTATTTAATACCACTCTCGCTCCATGAGGTCTCAGCC	540
Db	871	CATACCTCGCTGGTGAAGCTATTTAATACCACTCTCGCTCCATGAGGTCTCAGCC	930
QY	541	CAAAACCCCTACTAATCTTGGATGCTCCCTCCCTGCACTTTCAGGCCATACATTTCAATC	600
Db	931	CAAAACCCCTACTAATCTTGGATGCTCCCTCCCTGCACTTTCAGGCCATACATTTCAATC	990
QY	601	CCTGTTCTGCAACCAATGGAACAACTTCAGCAACAGAAATAAACACACTTCCGTTTACTA	660
Db	991	CCTGTTCTGCAACCAATGGAACAACTTCAGCAACAGAAATAAACACACTTCCGTTTACTA	1050
QY	661	GGACCTCTGTTTCCAAATCTGGAATAACCAATACCTCAAAACCTCACCTGTGTAAATTT	720
Db	1051	GGACCTCTGTTTCCAAATCTGGAATAACCAATACCTCAAAACCTCACCTGTGTAAATTT	1110
QY	721	AGCAATACATAGACACAAACAGCTCCCAATGCATCAGTGGGTAAACCTCCCAACAGA	780
Db	1111	AGCAATACATAGACACAAACAGCTCCCAATGCATCAGTGGGTAAACCTCCCAACAGA	1170
QY	781	ATAGTCTGCCTACCTCAGGAATAATTTTGTCTGTGTACCTCAGCTTATCATTTGTTG	840
Db	1171	ATAGTCTGCCTACCTCAGGAATAATTTTGTCTGTGTACCTCAGCTTATCATTTGTTG	1230
QY	841	AATGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTTATGACCATCTAC	900
Db	1231	AATGCTCTTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGCCCTTATGACCATCTAC	1290
QY	901	ACTGAACAGATTTTATACAAATCATCTGCTTACCTAAGCCCAACAAAGAGTACCCATT	960
Db	1291	ACTGAACAGATTTTATACAGTTATGTCAATCTAAGCCCGCAACAAAGAGTACCCATT	1350
QY	961	CTTCTTTTGTATCAGAGCAGGAGTGTAGGCAGACTAGGTACTGGCAATGGCAGTATC	1020
Db	1351	CTTCTTTTGTATCAGAGCAGGAGTGTAGGTGCACTAGGTACTGGCAATGGCAGTATC	1410
QY	1021	ACAACTCTACTCAGTTCTACAAACTATCTCAAGAAATAAATGTTGACATGGAACAG	1080
Db	1411	ACAACTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGTTGACATGGAACAG	1470
QY	1081	GTCACTGACTCCCTCGTCACTTGCAGAGTCAACTTAACTCCTCCTAGCAGCAGTAGTCTT	1140
Db	1471	GTGCGGAGTCCCTCGTCACTTGCAGAGTCAACTTAACTCCTCCTAGCAGCAGTAGTCTT	1530
QY	1141	CAAAATCGAAGAGCTTTTACACTTGTAAACCGCAAAAGAGGGGAACTGTGTTATTTTAA	1200
Db	1531	CAAAATCGAAGAGCTTTTACACTTGTAAACCGCTGAAAGAGGGGAACTGTGTTATTTTAA	1590
QY	1201	GGAGAAGAACGCTGTTATTTATGTTTAACTCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA	1260

Db 1591 GGGGAAGAATGCTGTATTATTAATCAATCCGAATCGTCACTGAGAAGATTAAAGAA 1650
Qy 1261 ATTGAGATCGAATCAATAGTAGAGAGAGAGCTTCAAAACACCGAAGCTGGGGCCTC 1320
Db 1651 ATTGAGATCGAATCAATAGTAGAGAGAGAGCTTCAAAACACCGAAGCTGGGGCCTC 1710
Qy 1321 CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1380
Db 1711 CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1770
Qy 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTTCCAGAAAT 1440
Db 1771 CTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTTCCAGAAAT 1830
Qy 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAATGGAAACCCCA 1481
Db 1831 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGAGTCCAA 1871

RESULT 36
AX007999
LOCUS AX007999 2055 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 22 from Patent WO9967395.
ACCESSION AX007999
VERSION AX007999.1 GI:9995696
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Perin, J.P., Rieger, F. and Alliel, P.M.
TITLE Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL Patent: WO 9967395-A 22 29-DEC-1999;
INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER
FRANCOIS (FR); ALLIEL PATRICK M (FR)
FEATURES
source Location/Qualifiers
1..2055
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS <1..87
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07392.1"
/db_xref="GI:9995697"
/translation="PKTANLVADITSLAKYQVVLKTLQGTYP"
ORIGIN

Query Match 90.1%; Score 1333.8; DB 6; Length 2055;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCACTATTTCTCTTACTGTCTCTTACCCCTTTTCGCTCTCACT 60
Db 391 ATGGCCCTCCCTTATCAATATTTTCTCTTACTGTCTCTTACTCCCTCTTCACTCTCACT 450
Qy 61 GCACCCCTCCATGCTGCTGACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 451 GCACCCCTCCATGCCGCTGATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGGAGA 510
Qy 121 ACGGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTCTAAGGGAATCTCC 180
Db 511 ATGACAGCTCCCGGAAATATTGATGCCCATCATATAGGAGTTTCTAAGGGAATCTCC 570
Qy 181 ACCTTCACTGCCCACCATATGCCCCCAACTGCTATTAATCTCTGCACTCTTTGCAATG 240
Db 571 ACCTTCACTGCCCACCATATGCCCCCAACTGCTATTAATCTCTGCACTCTTTGCAATG 630
Qy 241 CATGCAAACTCAATATTGACAGAGGAAATGATTAATCTCTAGTGTCTGAGGACTT 300
Db 631 CATGCAAACTCAATATTGACAGAGGAAATGATTAATCTCTAGTGTCTGAGGACTT 690

Qy 301 EGAGCCACTGTCTGTGTGACCTTACTTACCCCATACAGATATCTCTGATGGGGGTGAATTT 360
Db 691 GGAGTCACTGTCTGTGTGACCTTACTTACCCCATACAGTATCTCTGATGGGGGTGAGTT 750
Qy 361 GAAGTCAAGCAAG 420
Db 751 CAAGATCAGGCAAG 810
Qy 421 AGCACCCCTAGCCCTTACAAAGGAGTGTCTCTCAAAACTACATGAAACCCCTCCGTACC 480
Db 811 GGCACCTCTAGCCCTTACAAAGGAGTGTCTCTCAAAACTACATGAAACCCCTCCGTACC 870
Qy 481 CATACTCGCTCGGTGAGCCTATTTAATACCACTCCTGCGGTCTCATGAGTCTCAGCC 540
Db 871 CATACTCGCTCGGTGAGCCTATTTAATACCACTCCTGCGGTCTCATGAGTCTCAGCC 930
Qy 541 CAAAACCTTACTAACTGTGTGGATGTCCTCCCTCGCTTCCAGTCCATATTTCAATC 600
Db 931 CAAAACCTTACTAACTGTGTGGATGTCCTCCCTCGCTTCCAGTCCATATGTTTCAATC 990
Qy 601 CTGTTCCTGAAACAATGGAAACAATCTCAGCAGAGAAATAAACACCACTTCCGTTTAGTA 660
Db 991 CTGTTCCTGAAACAATGGAAACAATCTCAGCAGAGAAATAAACACCACTTCCGTTTAGTA 1050
Qy 661 GGAACCTTGTGTTTCCAACTCTGGAATAAACCACTCCTCAAACTCCTGTTGTAATTTT 720
Db 1051 GGAACCTTGTGTTTCCAACTCTGGAATAAACCACTCCTCAAACTCCTGTTGTAATTTT 1110
Qy 721 AGCAATACTATAGACACACACAGCTCCCAATGATGAGTGGGTGATACCTCCACACGA 780
Db 1111 AGCAATACTATAGACACACACAGCTCCCAATGATGAGTGGGTGATACCTCCACACGA 1170
Qy 781 ATAGTCTGCTACCTCCTCAGCAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
Db 1171 ATAGTCTGCTACCTCCTCAGCAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 1230
Qy 841 AATGGCTCTTCAGAACTCTATGTGCTCTCTCAATCTTATGTCCTCCCTATGACCATCTAC 900
Db 1231 AATGGCTCTTCAGAACTCTATGTGCTCTCTCAATCTTATGTCCTCCCTATGACCATCTAC 1290
Qy 901 ACTGAACAAGATTATATCAATCATGTCTGCTACCTAAGCCCCACAAACAAGAGTACCCATT 960
Db 1291 ACTGAACAAGATTATATCAAGTTATGTCTATGTCATATCTAAGCCCCACAAACAAGAGTACCCATT 1350
Qy 961 CTTCTCTTTGTTATCAGAGCAGAGTGTAGGAGAGTGTAGTGTGCTAGTGTGCGGTATC 1020
Db 1351 CTTCTCTTTGTTATCAGAGCAGAGTGTAGTGTGCTAGTGTGCTAGTGTGCGGTATC 1410
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 1411 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1470
Qy 1081 GTCATGATCTCCCTGTCACCTTGCAAGATCACTTAACCTCCCTAGCAGAGTAGTCCCTT 1140
Db 1471 GTGCGGACCTCCCTGTCCTGCTGCAAGTCACTTAACCTCCCTAGCAGAGTAGTCCCTT 1530
Qy 1141 CAAATCGAAGAGCTTTAGAGCTTGTAAACCGCAAAAGAGGGGAACTGTTTATTTTA 1200
Db 1531 CAAATCGAAGAGCTTTAGAGCTTGTAAACCGCTGAAAGAGGGGAACTGTTTATTTTA 1590
Qy 1201 GGAGAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATGTCATGAGAAAGTTAAAGAA 1260
Db 1591 GGAGAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATGTCATGAGAAAGTTAAAGAA 1650
Qy 1261 ATTGAGATCGAATCAATGATAGCAGAGAGAGCTTCAAAACACCGAAGCTGGGGCCTC 1320
Db 1651 ATTGAGATCGAATCAAAACGATAGCAGAGAGAGCTTCAAAACACCGAAGCTGGGGCCTC 1710
Qy 1321 CTGAGCAATGAGTGCCTGGGTTCTCCCTCTTCTAGGAGCTCTAGCAGCTCTAATATTG 1380
Db 1711 CTGAGCAATGAGTGCCTGGGTTCTCCCTCTTCTAGGAGCTCTAGCAGCTCTAATATTG 1770

```
QY 1381 TTACTCCTCTTTGGACCCCTGATCTTTAAACCTCTGTGTTAGTTGTTCTCTCCAGAAATT 1440
Db 1771 CTACTCCTCTTTGGACCCCTGATCTTTAAACCTCTGTGTTAACTTTGTCTCTCCAGAAATC 1830

QY 1441 GAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA 1481
Db 1831 GAAGCTGTAAACTACAAATGAGCCCAAGATGCAATCCAA 1871

RESULT 37
BD221806
LOCUS
DEFINITION
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses.
ACCESSION
BD221806
VERSION
BD221806.1 GI:33031576
KEYWORDS
JP 2002518051-A/1.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2599)
AUTHORS
Alliel,P.M., Perin,J.P. and Rieger,F.
TITLE
Nucleic sequence and deduced protein sequence family with human
endogenous retroviral motifs, and their uses
JOURNAL
Patent: JP 2002518051-A 1 25-JUN-2002;
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
COMMENT
OS Homo sapiens (human)
PN JP 2002518051-A/1
PD 25-JUN-2002
PP 23-JUN-1999 JP 2000556036
PR 23-JUN-1998 FR 98/07920
PI PATRICK M ALLIEL, JEAN PIERRE PERIN, FRANCOIS RIEGER PC
C12N15/09,A01K67/027,A61K31/711,A61K39/21,A61K48/00,A61P21/00, PC
A61P25/00,
PC A61P37/06,C07K14/15,C12Q1/70,C12Q1/68,C12N15/00 CC Nucleic
sequence and deduced protein sequence family with CC
human endogenous
CC retroviral motifs, and their uses
FH Key Location/Qualifiers
FT source 1..2599
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..2599
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 90.1%; Score 1333.8; DB 6; Length 2599;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTTTACTTGTCTCTTACCCCTTTTGGCTCTCACT 60
Db 929 ATGGCCCTCCCTTATCATATTTTCTTTTACTTGTCTTTTACCCCTTTTCACTCTCACT 988

QY 61 GCACCCCTCCATGCTGTGTACAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
Db 989 GCACCCCTCCATGCGCGTGTATGACAGTAGCTCCCTTTACCAAGAGTTTCTATGGAGA 1048

QY 121 ACGCGCTTCCCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
Db 1049 ATGACGCGTCCCGGAATATTGATGCCCATCATGTATAGGAGTTTCTTAAGGGAACCCCC 1108

QY 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCATG 240
Db 1109 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCATG 1168

QY 241 CATGCAAACTACTATTATGGACAGGGAAATGATTAATCCCTAGTTGTCCTGGAGCACTT 300
Db 1169 CATGCAAACTACTATTATGGACAGGGAAATGATTAATCCCTAGTTGTCCTGGAGCACTT 1228
```

```
QY 301 GGACCCACTGCTGTGTGGACTTACTTCCACCATAACAGTATGTCTGATGGGGTGGAAATT 360
Db 1229 GGAGTCACTGCTGTGTGGACTTACTTCCACCATAACAGTATGTCTGATGGGGTGGAGTT 1288

QY 361 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAAACCTGACCCGGGGACAT 420
Db 1289 CAAGATCAGGCACAGAGAAAAACAATGTAAAGGAAGTAAATCTCCCAAACCTCACCCGGGTACAT 1348

QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACCTACATGAACACCTCCGTACC 480
Db 1349 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAACCTACATGAACACCTCCGTACC 1408

QY 481 CATACTCGCTGTGTGAGCCTATTTAATACCAACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 1409 CATACTCGCTGTGTGAGCCTATTTAATACCAACCCCTCACTCGGCTCCATGAGGTCTCGGCC 1468

QY 541 CAAAACCTACTAACTGTGTGGATGTGCTCCCTCCCTGCACCTTCAGGCCCATCATTTCAATC 600
Db 1469 CAAAACCTACTAACTGTGTGGATGTGCTCCCTCCCTGAACTTTCAGGCCCATATGTTTCAATC 1528

QY 601 CCTGTTCTGAACAAATGGAACAACTTCAGCACAGAAATAAACACACACTTCCGTTTGTAGTA 660
Db 1529 CCTGTACTGAACAAATGGAACAACTTCAGCACAGAAATAAACACACACTTCCGTTTGTAGTA 1588

QY 661 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACCTGTGTGTAATAATT 720
Db 1589 GGACCTCTTGTTCCTCAATCTGGAATAAACCCATACCTCAAACTCACCTGTGTGTAATAATT 1648

QY 721 AGCAATACTATAGACACAAACCACTCCCAATGCAATGCAATGCAATGCAATGCAATGCAAT 780
Db 1649 AGCAATACTATAGACACAAACCACTCCCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1708

QY 781 ATAGTCTGCTTACCTCAGGAAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 840
Db 1709 ATAGTCTGCTTACCTCAGGAAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1768

QY 841 AATGCTCTTTCAGAAATCTATGTGCTTCTCTCAATTTTAGTGCCCTCTATGACCATCTAC 900
Db 1769 AATGCTCTTTCAGAAATCTATGTGCTTCTCTCAATTTTAGTGCCCTCTATGACCATCTAC 1828

QY 901 ACTGAACAAGATTTTATACAATCATGTGCTACTTAAGCCCAACCAAAAGAGTACCCATT 960
Db 1829 ACTGAACAAGATTTTATACAGTTATGTCTATATCTAAGCCCAACCAAAAGAGTACCCATT 1888

QY 961 CTTCCCTTTTGTATCAGAGCAGGAGTGTCTAGCAGACTAGTACTTGGCAATTTGGCAGTATC 1020
Db 1889 CTTCCCTTTTGTATAGGAGCAGGAGTGTCTAGTGCACCTAGTACTTGGCAATTTGGCAGTATC 1948

QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTTGACATGGAACAG 1080
Db 1949 ACNACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTTGGAACAG 2008

QY 1081 GTCAGTACTGCTGCTGTCACCTTGCAGATCAACTTAACCTCCCTAGCAGCAGTACTGCTTT 1140
Db 2009 GTGCGCAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2068

QY 1141 CAAATCGAAGAGCTTTTACACTTGTAAACCGCAAAAGAGGGGAACTCTGTTTATTTTATA 1200
Db 2069 CAAATCGAAGAGCTTTTACACTTGTAAACCGCTGAAGAGGGGAACTCTGTTTATTTTATA 2128

QY 1201 GGAGAAAGACCGTGTATTATTTAATCAATCCAGAAATGTCTACAGAAAGTTAAAGNA 1260
Db 2129 GGGGAAGAAATGCTGTTATTATTTAATCAATCCAGAAATGTCTACAGAAAGTTAAAGNA 2188

QY 1261 ATTCCAGATCGAATCAATGTAGACAGAGGAGCTTCAAAACACCGAGCTGGGGCCTC 1320
Db 2189 ATTCCAGATCGAATCAACAGTATAGCAGAGGAGCTTTCGAAACACTTGGACCCCTGGGGCCTC 2248

QY 1321 CTCAGCAATGATCCCTGGGTTCTCCCTCTTTTAGGACCTCTTAGCAGCTCTAATATTG 1380
Db 2249 CTCAGCAATGATGCTGGATCTCCCTCTTTTAGGACCTCTTAGCAGCTCTAATATTG 2308

QY 1381 TTACTCCTCTTTGGACCCCTGATATCTTTAAACCTCTTGTAAAGTTGTCCTCTCCAGAAATT 1440
```

Db 2309 CTACTCTCTTTGGACCTGTAATCTTAACTCTCTTGTAACTTTGTCCTCCAGAAATC 2368

Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAATGGAACCCCA 1481

Db 2369 GAAGCTGTAAGCTACAGATGGTCTTACAATGGAACCCCA 2409

RESULT 38

AX007978 2599 bp DNA linear PAT 06-SEP-2000

LOCUS Sequence 1 from Patent WO9967395.

DEFINITION AX007978

ACCESSION AX007978.1 GI:9995675

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Perin,J.P., Rieger,F. and Alliel,P.M.

AUTHORS Nucleic sequence and deduced protein sequence family with human

TITLE endogenous retroviral motifs, and their uses

JOURNAL Patent: WO 9967395-A 1 29-DEC-1999;

INST NAT SANTE RECH MED (FR); PERIN JEAN PIERRE (FR); RIEGER FRANCOIS (FR); ALLIEL PATRICK M (FR)

FEATURES

Location/Qualifiers

1..2599

source /organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 90.1%; Score 1333.8; DB 6; Length 2599;

Best Local Similarity 93.8%; Pred. No. 0;

Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTATCATCTTTCTTTACTGTTCTTACCCCTTTGGCTTCACCT 60

Db 929 ATGGCCCTCCCTATCATCTTTCTTTACTGTTCTTTTACCCCTTTTCACTCTCACT 988

Qy 61 GCACCCCTCCATCTGCTGACAAACAGTAGCTCCCTTACCAGAGTTTCTATGAAGA 120

Db 989 GCACCCCTCCATGCGCTGATGACAGTAGCTCCCTTACCAGAGTTTCTATGAAGA 1048

Qy 121 ACGGGCTTCTCGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAAATCC 180

Db 1049 ATGACGGCTCCGGAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAAATCC 1108

Qy 181 ACCTTCACTGCCACACCATATATGCGCGCACTGCTATTAACCTCTGCACTCTTTGCATG 240

Db 1109 ACCTTCACTGCCACACCATATATGCGCGCACTGCTATTAACCTCTGCACTCTTTGCATG 1168

Qy 241 CATGCAATATCTATTTGACAGAGGAATGATTAATCTAGTGTCTGAGGAGACTT 300

Db 1169 CATGCAATATCTATTTGACAGAGGAATGATTAATCTAGTGTCTGAGGAGACTT 1228

Qy 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACCATATGCTGTGATGGGGTGGAAAT 360

Db 1229 GGAGTCACTGTCTGTGGACTTACTTCAACCATACCATATGCTGTGATGGGGTGGAGTT 1288

Qy 361 CAAGGTGAGGCAAGAGAAAACAAGTAAGGAAGCAATCTCCCAATGACCCCGGGGACAT 420

Db 1289 CAAGTCAAGGCAAGAGAAAACAAGTAAGGAAGTAATCTCCCAATGACCCCGGGTACAT 1348

Qy 421 AGCACCCCTAGCCCTCAAGAGGACTAGTCTCTCAAACTACATGAACCCCTCCGTACC 480

Db 1349 GGCACCTCTAGCCCTCAAGAGGACTAGTCTCTCAAACTACATGAACCCCTCCGTACC 1408

Qy 481 CATACTCGCCCTGGTGGAGCTATTTTAATACCAACCCCTCACTCGGCTCCATGAGTCTCGGCC 540

Db 1409 CATACTCGCCCTGGTGGAGCTATTTTAATACCAACCCCTCACTCGGCTCCATGAGTCTCGGCC 1468

Qy 541 CAAAACCCCTACTAACTGTTGGATGCTCCCTGCACTTCAGGCCATACATTTCAATC 600

Db 1469 CAAAACCCCTACTAACTGTTGGATGCTCCCTGCACTTCAGGCCATATGTTTCAATC 1528

Qy 601 CTTGTTCTTGAACATGGAACAACTTCAGCAGAGAATAAACAACACTTCCTGTTTAGTA 660

Db 1529 CTTGTTCTTGAACATGGAACAACTTCAGCAGAGAATAAACAACACTTCCTGTTTAGTA 1588

Qy 661 GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCTCACTGTGTAAATTT 720

Db 1589 GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCTCACTGTGTAAATTT 1648

Qy 721 AGCAATATATAGACAAACACAGCTCCCAATGATGATGAGTGGTAAACACTCCCAACGA 780

Db 1649 AGCAATATATAGACAAACACAGCTCCCAATGATGATGAGTGGTAAACACTCCCAACGA 1708

Qy 781 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTGCTGTGTACCTCAGCCCTATCATCTTTG 840

Db 1709 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTGCTGTGTACCTCAGCCCTATCATCTTTG 1768

Qy 841 AATGGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900

Db 1769 AATGGCTCTTCAAGATCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1828

Qy 901 ACTGAAACAAGATTTATACAAATCATGTCGTAACCTTAAGCCCAACAAACAAAGAGTACCATT 960

Db 1829 ACTGAAACAAGATTTATACAGTTATGTCATATCTAAGCCCAACAAACAAAGAGTACCATT 1888

Qy 961 CTTCTCTTTTGTATACAGAGAGAGTGTAGGAGACTAGTACTGGCATTGGCAGTATC 1020

Db 1889 CTTCTCTTTTGTATACAGAGAGAGTGTAGGAGACTAGTACTGGCATTGGCAGTATC 1948

Qy 1021 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAATAAATGTCATCTGCAATGGAACAG 1080

Db 1949 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAATAAATGTCATCTGCAATGGAACAG 2008

Qy 1081 GTCACCTGACTCCCTGCTGCTTGTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTTT 1140

Db 2009 GTCGCGCATCTCCCTGCTGCTTGTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTTT 2068

Qy 1141 CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAGAGGGGGAACCTGTTTATTTTA 1200

Db 2069 CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAGAGGGGGAACCTGTTTATTTTA 2128

Qy 1201 GGAAGAGAGCTGTTTATTTATGTAATCAATCAGAAATGTCATCTGAGAAAGTTAAAGAA 1260

Db 2129 GGAAGAGAGCTGTTTATTTATGTAATCAATCAGAAATGTCATCTGAGAAAGTTAAAGAA 2188

Qy 1261 ATTGAGATCGAATACAATGATAGAGAGAGGAGCTTCAAAACACCGAAACCTGGGCTC 1320

Db 2189 ATTGAGATCGAATACAATGATAGAGAGAGGAGCTTCAAAACACCTGGGCTC 2248

Qy 1321 CTCAGCAATGAGTCCCTGGGTTCTCCCTTTTAGGACCTCTAGCAGCTCTAATATTG 1380

Db 2249 CTCAGCAATGAGTCCCTGGGTTCTCCCTTTTAGGACCTCTAGCAGCTCTAATATTG 2308

Qy 1381 TTAATCTCTTTGGACCTGATCTTTTAACTCTCTTAAAGTTTGTCTCTCCAGAAAT 1440

Db 2309 CTACTCTCTTTGGACCTGATCTTTTAACTCTCTTGTGTTAACTTTGTCTCTCCAGAAATC 2368

Qy 1441 GAAGCTGTTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481

Db 2369 GAAGCTGTTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 2409

RESULT 39

AF520481S2

LOCUS Homo sapiens individual 147 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.

DEFINITION AF520482

ACCESSION AF520482

VERSION AF520482.1

KEYWORDS GI:33410932

Db 1604 CTCAGCAATGATGCGCTGATTCCTCCCTCTTAGGACCTCTAGCAGCTATAATATTG 1663
 QY 1381 TTACTCTCTTTGGACCTGATCTTAACTCTCTTAAAGTTTCTCTCCAGCAATT 1440
 Db 1664 CTACTCTCTTTGGACCTGATCTTAACTCTCTTAAAGTTTCTCTCCAGCAATT 1723
 QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 Db 1724 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1764

RESULT 40
 AF520483S2 2694 bp DNA linear PRI 11-FEB-2004
 LOCUS Homo sapiens individual 147 allele B, envelope glycoprotein gene,
 DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
 ACCESSION AF520484
 VERSION AF520484.1 GI:33410936
 KEYWORDS 2 of 2
 SEGMENT Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
 Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
 TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
 in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBMED 14757826
 REFERENCE 2 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O. and Oriol,G.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
 d'Italie, Lyon 69364 cedex 07, France

FEATURES
 source
 1..2694
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="individual 147 allele B"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-q22"
 /sex="female"
 /cell_type="PBMC's"
 /note="endogenous_virus: HERV-W"
 misc_feature
 43..57
 /note="splice acceptor site"
 CDS
 284..1900
 /note="syncytin"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="AA017564.1"
 /db_xref="GI:33410937"
 /translation="MALPYHIFLFTVLLPSFTLTAPPPCRMTSSSPYQBFELRMQRP
 GNIDAPSYRLSKGTPTTAHMPNRCYHATLCMHANTHYWTGKMNPSCPGLGV
 TVCWYQTQSGMDGGVQDAEKHKEVIQSOLTRVHGTSSPYKGLDLKSLHETLR
 HTRVLNLETLTGLHVSQNPNTNWCILPLNRPVYSIPVPEOWNFSEINTTSV
 LVGLPSNLEITHSNLTCKVFSTNTVTSQICRWVTPTQIVCLPSGIFPFCVGSIA
 YRLUNGSESCFLUSFLVPWMTIYEQDLYSYVSKPRKRPVLPFVIGAVILGALG
 TGIGGITSTOYFKYKLSQELNGDMERVADSLVTIQDQNSLAALVQNRRLDLTLAE
 RGCTCLFGECCYVQSGIVTEKVEIRDIQRRAEELRNTGFWGLLSQWMPWILP
 FLGLAAIILLFLPGCTFNLLVNFVSRIEAVKLQMEPKVQSKTIYRPLDRPASP
 RSDYNDIKGTPEEISAAQPLLRNSAGSS"
 1705..1720
 misc_feature
 /note="splice acceptor site"
 LTR
 1959..2694

Query Match 90.1%; Score 1333.8; DB 9; Length 2694;
 Best Local Similarity 93.8%; Pred. No. 0;

Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATCATACACTTTTCTTTTACTGTCTTACCCTTTTCCCTCTCCT 60
 Db 284 ATGGCCCTCCCTTATCATATTTTCTTTTACTGTCTTACCCTTTTCCCTCTCCT 343
 QY 61 GCACCCCTCCATGCTGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
 Db 344 GCACCCCTCCATGCTGCTGTATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 403
 QY 121 ACGCCGCTTCTTGGAAATATTGATGCTCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
 Db 404 ATGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTCTTAAGGAAACCC 463
 QY 181 ACCTTCACCTGCCACACACCCATATGCCCCGCAACTGCTATAACTTGCCTACTTTTCATG 240
 Db 464 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATACTCTGCTGCTACTTTTCATG 523
 QY 241 CATGCAAAATCTCATTTATTGGACAGGAAATGATTAACTCTAGTTGTCTGAGGACTTT 300
 Db 524 CATGCAAAATCTCATTTATTGGACAGGAAATGATTAACTCTAGTTGTCTGAGGACTTT 583
 QY 301 GGAGCCACTGTCTGTGGACTTACTTACCCCATACAGTATGCTGATGGGGTGGAAATT 360
 Db 584 GGAGTCACTGTCTGTGGACTTACTTACCCCATACAGTATGCTGATGGGGTGGAGTT 643
 QY 361 CAAGGTCAAGCAAGGAAACAAAGTAAAGAGCAATCTCCCAACTGACCCCGGGGACAT 420
 Db 644 CAAGATCAGCAAGGAAACAAAGTAAAGAGTAACTCTCCCAACTGACCCCGGGTACAT 703
 QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTTCCGTACC 480
 Db 704 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTTCCGTACC 763
 QY 481 CATACTCGCTGTGTAGCCCTTATTAATACACCTTCTCGGTCCATGAGGTCTCAGCC 540
 Db 764 CATACTCGCTGTGTAGCCCTTATTAATACACCTTCTCGGTCCATGAGGTCTCAGCC 823
 QY 541 CAAACCCCTACTAACTGTGGTGTGCTCCCTCGCTTCTCAGGGCCATACATTTCAATC 600
 Db 824 CAAACCCCTACTAACTGTGGTGTGCTCCCTCGCTTCTCAGGGCCATATGTTCAATC 883
 QY 601 CTTGTTCTTGAACAATGGAACAACTTACGACAGAAATAAACACCACTTCCGTTTGTAGTA 660
 Db 884 CTTGTTCTTGAACAATGGAACAACTTACGACAGAAATAAACACCACTTCCGTTTGTAGTA 943
 QY 661 GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAACCTCTGTTGTAATAATTT 720
 Db 944 GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAACCTCTGTTGTAATAATTT 1003
 QY 721 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACCTCCCAACACGA 780
 Db 1004 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAACTCTCCCAACAA 1063
 QY 781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCTATCATTTGTTG 840
 Db 1064 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCTATCATTTGTTG 1123
 QY 841 AATGGCTCTTCAAGAACTATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
 Db 1124 AATGGCTCTTCAAGAACTATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1183
 QY 901 ACTGAAACAGATTTATACATCATGTCGTACCTTAAGCCCAACAAAGAGTACCCATT 960
 Db 1184 ACTGAAACAGATTTATACATCATGTCGTACCTTAAGCCCAACAAAGAGTACCCATT 1243
 QY 961 CTTTCT 1020
 Db 1244 CTTTCT 1303
 QY 1021 ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAAACAG 1080
 Db 1304 ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGGTGTGATGGAAACAG 1363

Qy	1081	GTCACTGACTCCCTGGTCACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Db	1364	GTGCGGACTCCCTGGTCACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1423
Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGGAACCTGTTTATTTTAA	1200
Db	1424	CAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAAGAGGGGGGAACCTGTTTATTTTAA	1483
Qy	1201	GGAGAAGAACGCTGTTATTATGTTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA	1260
Db	1484	GGGAAGAATGCTGTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA	1543
Qy	1261	ATTCGAGATCGAATACAAATGTAGAGCAGAGGAGCTTCAAAACACCGAAGCTGGGGCCTC	1320
Db	1544	ATTCGAGATCGAATACAAATGTAGAGCAGAGGAGCTTCAAAACACCGAAGCTGGGGCCTC	1603
Qy	1321	CTCAGCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
Db	1604	CTCAGCCAATGGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTAATATTG	1663
Qy	1381	TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTTCCAGAAAT	1440
Db	1664	CTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTTCCAGAAAT	1723
Qy	1441	GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA	1481
Db	1724	GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCCAA	1764

Search completed: February 21, 2005, 11:52:28
Job time : 6367.91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:25:25 ; Search time 763.929 Seconds
(without alignments)
11476.370 Million cell updates/sec

Title: US-09-319-156b-9
Perfect score: 1481
Sequence: 1 agggccctccctatcatcac.....gtcttacaatggaacccca 1481

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s.*
- 2: Geneseqn1980s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	1481	2	AAV43217 Multiple
2	1481	100.0	1481	2	Aax29703 Clone C15
3	1481	100.0	1481	12	Adg14846 MSRV asso
4	1481	100.0	2030	3	Aa63826 Nucleotid
5	1461.8	98.7	1629	3	Aa96625 DNA encod
6	1450	97.9	1481	9	Adb84400 MSRV-1 as
7	1389.8	93.8	2074	6	Aad41225 Human EMB
8	1385.4	90.2	2930	6	Aad24195 Human syn
9	1335.4	90.2	2946	2	Aax77526 Human sec
10	1335.4	90.2	2946	3	Aaz59468 Human sec
11	1335.4	90.2	2946	10	Adc38776 Human cdn
12	1333.8	90.1	1617	5	Aah20070 HERV-W en
13	1333.8	90.1	2055	3	Abn97948 Human ret
14	1333.8	90.1	2599	3	Abn97927 Human ret
15	1333.8	90.1	8523	7	Ad330988 Human gen
16	1333.8	90.1	10499	3	Abn97929 Human ret
17	1333.8	90.1	56093	6	Abf161744 Colon ade
18	1330.6	89.8	2781	5	Aaf55630 Nucleotid
19	1324.2	89.4	7582	3	Aa59215 Human end
20	1322.6	89.3	7582	2	Aax25665 Complete

21	1308.2	88.3	2782	5	AAH20069
22	1306.6	88.2	2782	2	AAX25661
23	1306.6	88.2	2782	3	AAAS9211
24	1298.2	87.7	3464	10	ADe09587
25	1298.2	87.7	9502	10	ADf59718
26	1282	86.6	46340	3	ABN97978
27	1278.8	86.3	1799	3	ABN97931
28	1277.4	86.3	161334	11	ACN44334
29	1239	83.7	2784	3	ABN97930
30	1131.6	76.4	1894	4	AAI14608
31	1131.6	76.4	1894	4	ABA56337
32	1131.6	76.4	1894	4	AAI35980
33	1131.6	76.4	1894	4	ABA45822
34	1131.6	76.4	1894	4	ABA45822
35	1131.6	76.4	1894	4	ABA25978
36	1131.6	76.4	1894	4	AAK00018
37	1131.6	76.4	1894	4	AAK04516
38	1131.6	76.4	1894	4	ABS29670
39	1131.6	76.4	1894	5	AAI04422
40	1116.4	75.4	1948	6	ABS04589
41	1116.4	75.4	1948	3	AAAS9209
42	1048	70.8	6394	5	AAAS84210
43	716.6	48.4	792	4	AAI23803
44	716.6	48.4	792	4	ABa68919
45	716.6	48.4	792	4	AAI49115

ALIGNMENTS

RESULT 1
AAV43217
ID AAV43217 standard; cDNA; 1481 BP.

XX AC AAV43217;

XX DT 29-DEC-1998 (first entry)

XX DE Multiple sclerosis associated retrovirus fragment 5.

XX KW Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene; env gene; rheumatoid arthritis-associated virus; ss.

XX OS Multiple sclerosis associated retrovirus.

XX FH Key Location/Qualifiers
XX CDS 1..1479

XX FT /*tag= a
XX FT /product= "Encodes protein AAV71068"
XX FT /transl_except= (pos:115-117, appears to code for a stop codon)
XX FT /note= "CDS does not contain a stop codon"

XX WO9823755-A1.

XX PD 04-JUN-1998.

XX PF 26-NOV-1997; 97WO-IB001452.

XX PR 26-NOV-1996; 96US-00756429.

XX PA (INMR) BIO MERIEUX.

XX PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
XX PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;

XX DR WPI; 1998-322732/28.

XX P-PSDB; AAV71068.

XX FT New nucleic acid from retroviruses - useful for diagnosis, prevention and treatment of, e.g. multiple sclerosis.

XX PS Disclosure; Page 184-185; 286pp; English.

XX The present sequence represents a multiple sclerosis (MS) associated
 CC retrovirus (MSRV) genomic fragment used in the method of the invention.
 CC The invention provides complete or partial genomic sequences of the MSRV-
 CC 1 pol gene, gag gene and env gene, and polypeptides encoded by these
 CC genes. The invention also provides antibodies raised against the
 CC polypeptides. The genomic sequences, polypeptides and antibodies are also
 CC claimed useful for diagnosing infection by MS and rheumatoid arthritis-
 CC associated viruses, and also for prevention and treatment of infection
 CC with these viruses
 XX
 SQ Sequence 1481 BP; 412 A; 410 C; 261 G; 398 T; 0 U; 0 Other;

Query Match 100.0%; Score 1481; DB 2; Length 1481;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTCTCTTACTGTCTTCTACCCCTTTTCGCTCTCACT 60
 DB 1 ATGGCCCTCCCTTATCATACTTTCTCTTACTGTCTTCTACCCCTTTTCGCTCTCACT 60
 QY 61 GCACCCCTCCCTTATCATACTTTCTCTTACTGTCTTCTACCCCTTTTCGCTCTCACT 120
 DB 61 GCACCCCTCCCTTATCATACTTTCTCTTACTGTCTTCTACCCCTTTTCGCTCTCACT 120
 QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAACTCC 180
 DB 121 AGCGGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAACTCC 180
 QY 181 ACCTTCACCTGCCACACCATATGCCCCGCACTGCTATAACTCTGCGCACTCTTTGCATG 240
 DB 181 ACCTTCACCTGCCACACCATATGCCCCGCACTGCTATAACTCTGCGCACTCTTTGCATG 240
 QY 241 CATGCAAACTACTTATTTGGACAGGAAATGATTAATCTTAGTTGCTCGGAGGACTT 300
 DB 241 CATGCAAACTACTTATTTGGACAGGAAATGATTAATCTTAGTTGCTCGGAGGACTT 300
 QY 301 GGAGCCACTGTCTGTTGGAGTACTTACCCCATACCATGATGTTGATGGGGGTGGAAT 360
 DB 301 GGAGCCACTGTCTGTTGGAGTACTTACCCCATACCATGATGTTGATGGGGGTGGAAT 360
 QY 361 CAAAGCTCAGGCAAGCAAAAAGTAAGGAGCAATCTCCCACTGACCCGGGACAT 420
 DB 361 CAAAGCTCAGGCAAGCAAAAAGTAAGGAGCAATCTCCCACTGACCCGGGACAT 420
 QY 421 AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAAACCCCTCCGTACC 480
 DB 421 AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAAACCCCTCCGTACC 480
 QY 481 CATACCTGGCTGGTGGAGCTATTTAATACACCCCTCACTCGGCTCCATGAGGCTCAGCC 540
 DB 481 CATACCTGGCTGGTGGAGCTATTTAATACACCCCTCACTCGGCTCCATGAGGCTCAGCC 540
 QY 541 CAAAACCTACTAACTGTTGGATGTGCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
 DB 541 CAAAACCTACTAACTGTTGGATGTGCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
 QY 601 CTGTTCTCTGAAACAATGGAACAATTTTACAGACAGAAATAAAACACCACTTCCTGTTAGTA 660
 DB 601 CTGTTCTCTGAAACAATGGAACAATTTTACAGACAGAAATAAAACACCACTTCCTGTTAGTA 660
 QY 661 GGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAACCTCTGTTGAAATTT 720
 DB 661 GGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAACCTCTGTTGAAATTT 720
 QY 721 AGCAATACTATAGACACACACCTCCCAATGATCAGGTGGGTAAACACTCCACACCA 780
 DB 721 AGCAATACTATAGACACACACCTCCCAATGATCAGGTGGGTAAACACTCCACACCA 780
 QY 781 ATAGTCTGCTACCTCAGAAATATTTTGTCTGTTGGTACCTCAGCCTATCATTTGTTG 840
 DB 781 ATAGTCTGCTACCTCAGAAATATTTTGTCTGTTGGTACCTCAGCCTATCATTTGTTG 840

QY 841 AATGCTCTTCAGAACTATATGTGCTTCTCTCAATTTCTAGTGCCCTTATGACCATCTAC 900
 DB 841 AATGCTCTTCAGAACTATATGTGCTTCTCTCAATTTCTAGTGCCCTTATGACCATCTAC 900
 QY 901 ACTGAACAAGATTTATACAACTATGTCTAGTCTTAAAGCCCAACAACAAGAGTACCCATT 960
 DB 901 ACTGAACAAGATTTATACAACTATGTCTAGTCTTAAAGCCCAACAACAAGAGTACCCATT 960
 QY 961 CTTCTTTTGTATCAGACGAGGAGTCTAGGACAGTACTGCGCATTTGGCAGTATC 1020
 DB 961 CTTCTTTTGTATCAGACGAGGAGTCTAGGACAGTACTGCGCATTTGGCAGTATC 1020
 QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
 DB 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
 QY 1081 GTCACTGACTCCCTGGTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1140
 DB 1081 GTCACTGACTCCCTGGTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1140
 QY 1141 CAAATCGAAGAGCTTTTAGACTTGTAACTCAATCCAGAAATTTGCTGAGAAAGTTAAAGAA 1200
 DB 1141 CAAATCGAAGAGCTTTTAGACTTGTAACTCAATCCAGAAATTTGCTGAGAAAGTTAAAGAA 1200
 QY 1201 GGAGAAGAACTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1260
 DB 1201 GGAGAAGAACTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1260
 QY 1261 ATTGAGATCGAATACAAATGTAGACGAGAGCTTCAAAAACACCGAAGCTGGGGCTC 1320
 DB 1261 ATTGAGATCGAATACAAATGTAGACGAGAGCTTCAAAAACACCGAAGCTGGGGCTC 1320
 QY 1321 CTCGACCAATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 DB 1321 CTCGACCAATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 QY 1381 TTACTCTCTTTGGACCTCTTAACTCTTAACTCTTAACTCTTAACTCTTAACTCTTAACTCTTAACT 1440
 DB 1381 TTACTCTCTTTGGACCTCTTAACTCTTAACTCTTAACTCTTAACTCTTAACTCTTAACTCTTAACT 1440
 QY 1441 GAAGCTGTAAGCTACAGATGCTTCAAAATGGAACCCCA 1481
 DB 1441 GAAGCTGTAAGCTACAGATGCTTCAAAATGGAACCCCA 1481

RESULT 2
 AAX29703
 ID AAX29703 standard; DNA; 1481 BP.
 XX
 AC AAX29703;
 XX
 DT 17-OCT-2003 (revised)
 DT 27-AUG-2003 (revised)
 DT 08-JUN-1999 (first entry)
 XX
 DE Clone C15 from MSRV-1.
 XX
 KW Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
 KW rheumatoid polyarthritis; ss.
 XX
 OS Viruses.
 XX
 PN FR2765588-A1.
 XX
 PD 08-JAN-1999.
 XX
 PF 07-JUL-1997; 97FR-00008816.
 XX
 PR 07-JUL-1997; 97FR-00008816.
 XX
 PA (INMR) BIO MERIEUX.
 XX
 DR WPI; 1999-098275/09.

PI Perron H, Beseme F, Bedin F, Paranhos-Baccalla G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX WPI; 2004-032461/03.
XX New isolated nucleic acid and their fragments having the pol gene of a
PT retrovirus, useful for diagnosing, preventing and/or treating multiple
PT sclerosis and/or rheumatoid arthritis.
XX Disclosure; SEQ ID NO 105; 193pp; English.
XX The invention relates to an isolated nucleic acid which comprises the pol
CC gene of a retrovirus associated with multiple sclerosis or rheumatoid
CC arthritis. The methods and compositions of the present invention are
CC useful for diagnosing, preventing and/or treating multiple sclerosis
CC and/or rheumatoid arthritis. The present sequence is used in the
CC exemplification of the invention.
XX Sequence 1481 BP; 412 A; 410 C; 261 G; 398 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1481; DB 12; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTCTTACCCCTTTGCTCTCACT 60
DB 1 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTCTTACCCCTTTGCTCTCACT 60
QY 61 GCACCCCTCCCTGCTGTGTACACAGTAGTCTCCCTTTACCAAGAGTTTCTATGAAGA 120
DB 61 GCACCCCTCCCTGCTGTGTGTACACAGTAGTCTCCCTTTACCAAGAGTTTCTATGAAGA 120
QY 121 ACGCGCTTCTGGAAATATTGATGCCCATATATAGAGTTTATCTAAGGAACTCC 180
DB 121 ACGCGCTTCTGGAAATATTGATGCCCATATATAGAGTTTATCTAAGGAACTCC 180
QY 181 ACCTTCACCTGCCACACCCATATGCGCCGCACTGTATACTCTGCCACTCTTGCATG 240
DB 181 ACCTTCACCTGCCACACCCATATGCGCCGCACTGTATACTCTGCCACTCTTGCATG 240
QY 241 CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTTAGTTGTCTCGAGGACTT 300
DB 241 CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTTAGTTGTCTCGAGGACTT 300
QY 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCATTACAGTATGTCATGGGGTGAATTT 360
DB 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCATTACAGTATGTCATGGGGTGAATTT 360
QY 361 CRAAGTCAGGCAAGAGAAACAAAGTAAAGAGAGCAATCTCCAACTGACCCGGGACAT 420
DB 361 CRAAGTCAGGCAAGAGAAACAAAGTAAAGAGAGCAATCTCCAACTGACCCGGGACAT 420
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
DB 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
QY 481 CATACTGCGCTGTGGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB 481 CATACTGCGCTGTGGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
QY 541 CAAACCCCTACTAATCTGTGGATGCTCTCCCTGCACCTTCAGGCGCATACATTTCAATC 600
DB 541 CAAACCCCTACTAATCTGTGGATGCTCTCCCTGCACCTTCAGGCGCATACATTTCAATC 600
QY 601 CCTGTTCTGAACAATGGAACTTTCAGCACAGAAATAAACAACCTTCGTTTGTAGTA 660
DB 601 CCTGTTCTGAACAATGGAACTTTCAGCACAGAAATAAACAACCTTCGTTTGTAGTA 660
QY 661 GGACCTCTTGTTCCTCAATCTGGAAATTAACCCATACCTCAACCTCACTGTGTAAATTT 720
DB 661 GGACCTCTTGTTCCTCAATCTGGAAATTAACCCATACCTCAACCTCACTGTGTAAATTT 720
QY 721 AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCACACGA 780

DB 721 AGCAATACTATAGACACAAACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCACACGA 780
QY 781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTTACCTCAGCTATCATTTGTTG 840
DB 781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTTACCTCAGCTATCATTTGTTG 840
QY 841 AATGCTCTTTCAGAACTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
DB 841 AATGCTCTTTCAGAACTATGTGCTTCTCTCATTTCTAGTGGCCCTATGACCATCTAC 900
QY 901 ACTGAACAAGATTATTAACAATCATGTGCTAAAGCCCAACAAAGAGTAGTACCATT 960
DB 901 ACTGAACAAGATTATTAACAATCATGTGCTAAAGCCCAACAAAGAGTAGTACCATT 960
QY 961 CTTCTCTTTTGTATCAGAGCAGGAGTCTAGGACAGTGTAGTCTGGCATTGGCAGTATC 1020
DB 961 CTTCTCTTTTGTATCAGAGCAGGAGTCTAGGACAGTGTAGTCTGGCATTGGCAGTATC 1020
QY 1021 ACAACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGTGTGATCGAACAG 1080
DB 1021 ACAACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGTGTGATCGAACAG 1080
QY 1081 GTCACTGACTCCCTGGTCACTTGTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCT 1140
DB 1081 GTCACTGACTCCCTGGTCACTTGTCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCT 1140
QY 1141 CAAATCGAAGAGCTTTAGACTTGTAAACGCGCAAGAGAGGGGAACTGTTTATTTTA 1200
DB 1141 CAAATCGAAGAGCTTTAGACTTGTAAACGCGCAAGAGAGGGGAACTGTTTATTTTA 1200
QY 1201 GGAGAGAAACCTGTTTATTTATTAATCAATCCAGAAATTTGTCACCTGAGAAAGTTAAAGAA 1260
DB 1201 GGAGAGAAACCTGTTTATTTATTAATCAATCCAGAAATTTGTCACCTGAGAAAGTTAAAGAA 1260
QY 1261 ATTGCGAGATCGAATACAAATGTAGACGAGAGCTTCAAAACACCGAAGCGTGGGCGCTC 1320
DB 1261 ATTGCGAGATCGAATACAAATGTAGACGAGAGCTTCAAAACACCGAAGCGTGGGCGCTC 1320
QY 1321 CTCAGCAGTGGAGTCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 1321 CTCAGCAGTGGAGTCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTTGTAAAGTTGTCTCTTCCAGAAAT 1440
DB 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTTGTAAAGTTGTCTCTTCCAGAAAT 1440
QY 1441 GAAGCTGTAAAGCTACAGATGCTTCAAAATGGNAACCCCA 1481
DB 1441 GAAGCTGTAAAGCTACAGATGCTTCAAAATGGNAACCCCA 1481
RESULT 4
AAA63826
ID AAA63826 standard; DNA; 2030 BP.
XX
AC AAA63826;
XX 06-AUG-2003 (revised)
DT 04-DEC-2000 (first entry)
XX
DE Nucleotide sequence of the MSRV-1 3' env and LTR regions.
XX MSRV-1; pol region; long terminal repeat; LTR; RU5 region; retrovirus;
KW ss.
XX Multiple sclerosis associated retrovirus.
XX Location/Qualifiers
FH 1..1629
FT /tag= a
FT /note= "Contains one termination codon"
FT sig_peptide 1..81

FT CAAT_signal /*tag= b
FT 1800..1807
FT /*tag= c
FT CAAT_signal 1858..1864
FT /*tag= d
FT TATA_signal 1906..1911
FT /*tag= e
FT polyA_signal 1996..2002
FT /*tag= f
XX WO200047745-A1.
XX 17-AUG-2000.
XX 15-FEB-2000; 2000WO-18000159.
XX 15-FEB-1999; 99EP-00420041.
XX (INMR) BIO MERIEUX.
XX Paranhos-Baccala G, Perron H, Komurian-Pradel P;
XX WPI; 2000-506097/45.
XX P-PSDB; AAB08195.
XX Nucleotide fragment of LTR-RUS region from Multiple Sclerosis retrovirus
XX (MSRV) used to detect the presence of MSRV-1 retrovirus in a biological
XX sample.
XX Disclosure; Fig 2; 23pp; English.
XX The present sequence represents the nucleotide sequence corresponding to
XX the 3' env region and long terminal repeat sequences from clone Cl6 of
XX Multiple Sclerosis retrovirus (MSRV-1). The specification describes a
XX long terminal repeat (LTR)-RUS region which encodes the expression of a
XX MSRV-1 protein. This is unusual for LTRs, in particular in the RUS
XX region. The sequence includes CAAT and TATA signals which are present in
XX the U3 and R regions and are not directed towards the CDS indicated in
XX the features table. Probes and antibodies to the MSRV-1 retrovirus
XX protein and encoding polynucleotide sequences are used to detect the
XX presence of MSRV-1 retrovirus in a biological sample. (Updated on 06-AUG-
XX 2003 to correct OS field.)
XX
XX SQ Sequence 2030 BP; 574 A; 559 C; 387 G; 510 T; 0 U; 0 Other;
Query Match 100.0%; Score 1481; DB 3; Length 2030;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTCTCTTACCCCTTTGCGCTCTCACT 60
DB 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTCTCTTACCCCTTTGCGCTCTCACT 60
QY 61 GCACCCCTCCATCGTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
DB 61 GCACCCCTCCATCGTGTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
QY 121 ACGGGGCTTCTGGAAATATGTATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
DB 121 ACGGGGCTTCTGGAAATATGTATGCCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
QY 181 ACCTTCACTGCCACACCATATGCCCCGACCTGCTATTAACCTGCGCACTTTTGCATG 240
DB 181 ACCTTCACTGCCACACCATATGCCCCGACCTGCTATTAACCTGCGCACTTTTGCATG 240
QY 241 CATCAATATCTCAATTATGGACAGGAAATGATTAACTCTAGTGTCTTGGAGGACTT 300
DB 241 CATCAATATCTCAATTATGGACAGGAAATGATTAACTCTAGTGTCTTGGAGGACTT 300
QY 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGATGTCTGATGGGGTGGAAAT 360
DB 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGATGTCTGATGGGGTGGAAAT 360

QY 361 CAAGTCTAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
DB 361 CAAGTCTAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
DB 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
QY 481 CATACTCGCCTGGTGAAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
DB 481 CATACTCGCCTGGTGAAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
QY 541 CAAAACCCCTAATACTGTTGGATGCTCCCTCGCACTTCCAGGCCATACATTTCAATC 600
DB 541 CAAAACCCCTAATACTGTTGGATGCTCCCTCGCACTTCCAGGCCATACATTTCAATC 600
QY 601 CCTGTTCTGTAACAATGGAACTTACGACAGAAATAAACACCACTTCCGTTTTAGTA 660
DB 601 CCTGTTCTGTAACAATGGAACTTACGACAGAAATAAACACCACTTCCGTTTTAGTA 660
QY 661 GGACCTCTGTTTCCAACTCGGAAATAACCCATACCTCAAACTCACTGTGTAAATTT 720
DB 661 GGACCTCTGTTTCCAACTCGGAAATAACCCATACCTCAAACTCACTGTGTAAATTT 720
QY 721 AGCAATATATAGACAAACAGCTCCCAATGATCAGGTGGTAAACACCTCCACACGA 780
DB 721 AGCAATATATAGACAAACAGCTCCCAATGATCAGGTGGTAAACACCTCCACACGA 780
QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTGCTGGTACCTCAGCCTATCAATTTGTTG 840
DB 781 ATAGTCTGCTACCTCAGGAATATTTTGTGCTGGTACCTCAGCCTATCAATTTGTTG 840
QY 841 AATGGCTCTTCAAGATCTATGCTTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
DB 841 AATGGCTCTTCAAGATCTATGCTTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
QY 901 ACTGAACAAGATTTATACAAATCATGTGCTGTAACCTAAGCCCAACAAAGAGTACCCATT 960
DB 901 ACTGAACAAGATTTATACAAATCATGTGCTGTAACCTAAGCCCAACAAAGAGTACCCATT 960
QY 961 CTTCTCTTTTGTATCAGACGAGGAGTCTAGGAGACTAGGTACTGCGATTGCGAGTATC 1020
DB 961 CTTCTCTTTTGTATCAGACGAGGAGTCTAGGAGACTAGGTACTGCGATTGCGAGTATC 1020
QY 1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGACAG 1080
DB 1021 ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGACAG 1080
QY 1081 GTCACCTGACTCCCTGCTACCTTTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1140
DB 1081 GTCACCTGACTCCCTGCTACCTTTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1140
QY 1141 CAAAATCGAAGAGCTTTAGACTTTGCTAACCGCCAAAGAGGGGAACTGTTTATTTT 1200
DB 1141 CAAAATCGAAGAGCTTTAGACTTTGCTAACCGCCAAAGAGGGGAACTGTTTATTTT 1200
QY 1201 GGAGAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATTTCTCACTGAGAAAGTTAAAGAA 1260
DB 1201 GGAGAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATTTCTCACTGAGAAAGTTAAAGAA 1260
QY 1261 ATTTCGAGATCGAATAACAATGTAGACGAGAGGACTTCAAAACACCGCAACCGTGGGGCTC 1320
DB 1261 ATTTCGAGATCGAATAACAATGTAGACGAGAGGAGCTTCAAAACACCGCAACCGTGGGGCTC 1320
QY 1321 CTGAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1380
DB 1321 CTGAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTTAATATTG 1380
QY 1381 TTAATCTCTCTTTGGACCCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAAT 1440
DB 1381 TTAATCTCTCTTTGGACCCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAAT 1440
QY 1441 GAAGCTGTAAAGCTACAGATGGTCTTTACAAATGGAAACCCCA 1481

Db	1441	GAAGCTGTAAGACTACAGATGGTCTTACAAATGGAACCCCA	1481	
Db	1441	GAAGCTGTAAGACTACAGATGGTCTTACAAATGGAACCCCA	1481	
Qy	121	ACGGGCTTCTGGAAATATTGATGCCCATATATAGGAGTTTATCTTAAGGGAACCTCC	180	
Db	121	ACGGGCTTCTGGAAATATTGATGCCCATATATAGGAGTTTATCTTAAGGGAACCTCC	180	
Qy	181	ACCTTCACTGCCACACCCATATGCCCCGGAACCTGCTATAAATCTGCGCACTCTTTGCAATG	240	
Db	181	ACCTTCACTGCCACACCCATATGCCCCGGAACCTGCTATAAATCTGCGCACTCTTTGCAATG	240	
Qy	241	CATGCAAAATCTCATTTATTTGGACAGGGGAAATGATTATCTAGTTGTCCTGGAGGACTT	300	
Db	241	CATGCAAAATCTCATTTATTTGGACAGGGGAAATGATTATCTAGTTGTCCTGGAGGACTT	300	
Qy	301	GGAGCCACTGCTGTTGGACTTACTTTCACCCATACCCAGTATGTCTGATGCGGGTGGAAAT	360	
Db	301	GGAGCCACTGCTGTTGGACTTACTTTCACCCATACCCAGTATGTCTGATGCGGGTGGAAAT	360	
Qy	361	CAAGGTGAGGCAAGAGAAAAACAAGTAAAGAGCAATCTCCCACTACACCCGGGGACAT	420	
Db	361	CAAGGTGAGGCAAGAGAAAAACAAGTAAAGAGCAATCTCCCACTACACCCGGGGACAT	420	
Qy	421	AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAATACATAGAAACCCCTCGTACC	480	
Db	421	AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAATACATAGAAACCCCTCGTACC	480	
Qy	481	CATCTCGCTGCTGAGCCTATTTAATACCACTCTCTCGGCTCCATGAGTCTCAGCC	540	
Db	481	CATCTCGCTGCTGAGCCTATTTAATACCACTCTCTCGGCTCCATGAGTCTCAGCC	540	
Qy	541	CAAAACCTTACTAACTGTTGGATGCGCTCCCTCGCACTTCAGGCCCATAACATTTCAATC	600	
Db	541	CAAAACCTTACTAACTGTTGGATGCGCTCCCTCGCACTTCAGGCCCATAACATTTCAATC	600	
Qy	601	CCTGTTCTGAAACAATGGAAACAATCTCAGCACAGAAATAAACAACAATCTCTTTTAGTA	660	
Db	601	CCTGTTCTGAAACAATGGAAACAATCTCAGCACAGAAATAAACAACAATCTCTTTTAGTA	660	
Qy	661	GGACCTCTGTTTCCAACTCGAAATAACCCATACCTCAAACTCAGCTGCTGTAATAATTT	720	
Db	661	GGACCTCTGTTTCCAACTCGAAATAACCCATACCTCAAACTCAGCTGCTGTAATAATTT	720	
Qy	721	AGCAATACTATAGACAACAACAGCTCCCAATCATCATCAGGTGGGTAAACCTCCCAACA	780	
Db	721	AGCAATACTATAGACAACAACAGCTCCCAATCATCATCAGGTGGGTAAACCTCCCAACA	780	
Qy	781	ATAGTCTGCTACCTCAGGAATATTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	840	
Db	781	ATAGTCTGCTACCTCAGGAATATTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	840	
Qy	841	AATGGCTCTTCAAGATCTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900	
Db	841	AATGGCTCTTCAAGATCTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900	
Qy	901	ACTGAACAAGATTTATACAAATCATGTCGTAACCTGAGCCCAACAACAAAGAGTACCATT	960	
Db	901	ACTGAACAAGATTTATACAAATCATGTCGTAACCTGAGCCCAACAACAAAGAGTACCATT	960	
Qy	961	CTTCTCTTTGTTATCAGACGAGGAGTCTTAGCAGACTAGGTACTGSCATGTCGACTATC	1020	
Db	961	CTTCTCTTTGTTATCAGACGAGGAGTCTTAGCAGACTAGGTACTGSCATGTCGACTATC	1020	
Qy	1021	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTACATGGAACAG	1080	
Db	1021	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTACATGGAACAG	1080	
Qy	1081	GTCACTGACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140	
Db	1081	GTCACTGACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140	
Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCCAAGAGGGGGAAACCTGTTTATTTTA	1200	
Db	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCCAAGAGGGGGAAACCTGTTTATTTTA	1200	
Qy	1201	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1260	
Db	1201	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1260	
Qy	1261	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1320	
Db	1261	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1320	
Qy	1321	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1380	
Db	1321	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1380	
Qy	1381	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1440	
Db	1381	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1440	
Qy	1441	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1500	
Db	1441	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1500	
Qy	1501	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1560	
Db	1501	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1560	
Qy	1561	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1620	
Db	1561	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1620	
Qy	1621	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1680	
Db	1621	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1680	
Qy	1681	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1740	
Db	1681	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1740	
Qy	1741	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1800	
Db	1741	GGAGAAGAACGCTTTATTTATGTTAATCAATCAATCAATCAATCAATCAATCAATCAATCA	1800	
Qy	1801	GGAGAAGAAC		

Db 1201 GGAGAGAACGCTGTTATATGTTATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
Qy 1261 ATTGAGATCGAATACATAGTAGAGAGAGAGCTTCAAAACACCGAAGCTGGGCGCTC 1320
Db 1261 ATTGAGATCGAATACATAGTAGAGAGAGAGCTTCAAAACACCGAAGCTGGGCGCTC 1320
Qy 1321 CTCAGCAATGGATGCCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATG 1380
Db 1321 CTCAGCAATGGATGCCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATG 1380
Qy 1381 TTACTCCTCTTTGGACCTGTATCTTTAACTCTTTGTTAAGTTCTCTTCCAGAAAT 1440
Db 1381 TTACTCCTCTTTGGACCTGTATCTTTCAACTCTTTGTTAAGTTCTCTTCCAGAAAT 1440
Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 1441 GAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCCA 1481

RESULT 6
ADB84400
ID ADB84400 standard; DNA; 1481 BP.
AC ADB84400;
XX
DT 04-DEC-2003 (first entry)
XX
MSRV-1 associated DNA sequence #19.
DE ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
KW reverse transcriptase; ribonuclease H.
XX
XX Unidentified;
OS
PN US2003039664-A1.
XX
XX 27-FEB-2003.
PD
XX 26-NOV-1997; 97US-00979847.
PF
XX 26-NOV-1996; 96US-00756429.
PR
XX (PERR/) PERRON H.
PA (BESE/) BESEME F.
PA (BEDI/) BEDIN F.
PA (PARA/) PARANHOS-BACCALA G.
PA (KOMU/) KOMURIAN-PRADEL F.
PA (JOLI/) JOLIVET-REYNAUD C.
PA (MAND/) MANDRAND B.
PA (GARS/) GARSON J A.
PA (TUKE/) TUKE P W.
XX
XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX
XX WPI; 2003-512253/48.
DR
XX
XX New isolated or purified nucleic acid associated with multiple sclerosis
PT and/or rheumatoid arthritis, useful for detecting a virus associated with
PT multiple sclerosis or rheumatoid arthritis in a biological sample.
XX
XX Claim 31; Page 77-78; 193pp; English.
PS
XX The invention relates to an isolated or purified nucleic acid from a
XX virus associated with multiple sclerosis and/or rheumatoid arthritis,
CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
CC proteins or defined peptides (including immunodominant peptides,
CC antigenic peptides or conserved motifs). Also included are a process for
CC detecting a virus associated with multiple sclerosis or rheumatoid
CC arthritis in a biological sample, a nucleic acid probe for the detection
CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a

CC primer for the amplification by polymerisation of a nucleic acid of a
CC viral material associated with multiple sclerosis or rheumatoid
CC arthritis, a polypeptide exhibiting an inhibitory activity on the
CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
CC and an antibody directed against the MSRV-1 virus obtained by
CC immunologically reacting a human or animal body or cells with an
CC immunogenic agent consisting of the antigenic polypeptide defined above.
CC The nucleic acids are useful for detecting a biological sample a virus
CC associated with multiple sclerosis or rheumatoid arthritis, or for
CC detecting in a biological sample, the presence of or exposure to a virus
CC associated with multiple sclerosis or rheumatoid arthritis. The present
CC sequence is a claimed MSRV-associated sequence whose identity cannot be
CC accurately determined. Note: The SEQ ID numbers for the sequences as
CC displayed in the main body of the patent do not match the SEQ ID numbers
CC in the sequence listing. Consequently those sequences mentioned in the
CC claims may not be the sequences the authors intended to claim.
XX
SQ Sequence 1481 BP; 405 A; 403 C; 257 G; 385 T; 0 U; 31 Other;
Query Match 97.9%; Score 1450; DB 9; Length 1481;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1450; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
Qy 1 ATGGCCCTCCCTTATCATACACTTTTCTTTACTGTTCTTTACCCCTTTCCGCTCTCACT 60
Db 1 ATGGCCCTCCCTTATCATACACTTTTCTTTACTGTTCTTTACCCCTTTCCGCTCTCACT 60
Qy 61 GCACCCCTCCCTGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTTATGAAGA 120
Db 61 GCACCCCTCCCTGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTTATGAAGN 120
Qy 121 ACGCGCTTCTCTGGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
Db 121 ACGCGCTTCTCTGGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCN 180
Qy 181 ACCTTCACCTGCCACACCCATATATGATGCCCACTGCTATTAACCTCTCTTGCATN 240
Db 181 ACCTTCACCTGCCACACCCATATATGATGCCCACTGCTATTAACCTCTCTTGCATN 240
Qy 241 CATGCAATTAACCTATTTTGGACAGGAAATGATTAACTCTAGTTGTCTCGAGGACTT 300
Db 241 CATGCAATTAACCTATTTTGGACAGGAAATGATTAACTCTAGTTGTCTCGAGGACTN 300
Qy 301 GGAGCCACTGTCTGTTGGACTTTACTTACCCTACCAATACAGTATGTCTGATGGGTTGAATT 360
Db 301 GGAGCCACTGTCTGTTGGACTTTACTTACCCTACCAATACAGTATGTCTGATGGGTTGAATTN 360
Qy 361 CAAGGTGAGCAAGAGAAAAAACAAGTAAAGAGAGCAATCTCCCACTGACCCCGGACAT 420
Db 361 CAAGGTGAGCAAGAGAAAAAACAAGTAAAGAGAGCAATCTCCCACTGACCCCGGACAN 420
Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCTCCGTACC 480
Db 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCTCCGTACN 480
Qy 481 CATACTCGCTGTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 481 CATACTCGCTGTGAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCN 540
Qy 541 CAAACCCCTACTAATCTGTTGGATGTGCTCCCTCGACTTCAGGGCATACATTTCAATC 600
Db 541 CAAACCCCTACTAATCTGTTGGATGTGCTCCCTCGACTTCAGGGCATACATTTCAATN 600
Qy 601 CCTGTTCTGAAACAATGGAACTTTCAGCAGAGAAATAAACACCACTTCGTTTAGTN 660
Db 601 CCTGTTCTGAAACAATGGAACTTTCAGCAGAGAAATAAACACCACTTCGTTTAGTN 660
Qy 661 GGAACCTCTGTTTCCAAATCTGGAATAAACCCATACCTCAAACTCCTCTGTGTAATAATTT 720
Db 661 GGAACCTCTGTTTCCAAATCTGGAATAAACCCATACCTCAAACTCCTCTGTGTAATAATN 720
Qy 721 AGCAATACTATAGACACCAACCTCCATGATGAGTGGGTAAACCTCCACACGA 780
Db 721 AGCAATACTATAGACACCAACCTCCATGATGAGTGGGTAAACCTCCACACGA 780

Db 721 AGCAATATATAGACACACACAGCTCCCAATGCATCAGTGGGTAAACCTCCACACGN 780
Qy 781 ATAGTCTGCTACCTCCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
Db 781 ATAGTCTGCTACCTCCTCAGGAATATTTTTTGTCTGTGTACCTCAGCCTATCATTTGTTN 840
Qy 841 AATGGCTCTTCAGAACTATATGTCTTCTCTCATTTCTTAGTCCCTTATGACCATCTAC 900
Db 841 AATGGCTCTTCAGAACTATATGTCTTCTCTCATTTCTTAGTCCCTTATGACCATCTAN 900
Qy 901 ACTGAACAGATTTATACAACTATGTCTGCTACCTTAAGCCCAACAAAGAGTACCCATT 960
Db 901 ACTGAACAGATTTATACAACTATGTCTGCTACCTTAAGCCCAACAAAGAGTACCCATN 960
Qy 961 CTTCTCTTTTGTATCAGACGAGGTGCTAGGACAGTACCTAGGCAATGGCAATGCGATTC 1020
Db 961 CTTCTCTTTTGTATCAGACGAGGTGCTAGGACAGTACCTAGGCAATGGCAATGCGAGTANN 1020
Qy 1021 ACAACCTCTACTAGTCTTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 1021 ACAACCTCTACTAGTCTTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACNN 1080
Qy 1081 GTCACGTACTCCCTGCTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Db 1081 GTCACGTACTCCCTGCTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCNN 1140
Qy 1141 CAAATCGAAGAGCTTTAGACTGTCTAACCGCCCAAGAGGGGAACTTTTATTTTA 1200
Db 1141 CAAATCGAAGAGCTTTAGACTGTCTAACCGCCCAAGAGGGGAACTTTTATTTNN 1200
Qy 1201 GGAAGAGAGCGCTTATATGTTAATCAATCCAGAAATGTCACTCAGAAAGTTAAAGAA 1260
Db 1201 GGAAGAGAGCGCTTATATGTTAATCAATCCAGAAATGTCACTCAGAAAGTTAAAGNN 1260
Qy 1261 ATTGAGATCGAATCAATAGAGCAGAGGAGCTTCAAAACACCGAAGCTGGGCGCTC 1320
Db 1261 ATTGAGATCGAATCAATAGAGCAGAGGAGCTTCAAAACACCGAAGCTGGGCGCNN 1320
Qy 1321 CTCAGCAATGGAATGAGCTGGGTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 1321 CTCAGCAATGGAATGAGCTGGGTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTNN 1380
Qy 1381 TTACTCTCTTTGGACCTCTATCTTTAACTCTCTTAAAGTTTGTCTCTCCAGAAAT 1440
Db 1381 TTACTCTCTTTGGACCTCTATCTTTAACTCTCTTAAAGTTTGTCTCTCCAGAAANN 1440
Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481

RESULT 7
ID AAD41225 standard; cDNA; 2074 BP.
XX
AC AAD41225;
XX
DT 30-OCT-2002 (first entry)
XX
DE Human EMBRY-2 cDNA.
XX
KW Human; embryogenesis associated protein; AIDS; reproductive disorder;
KW infertility; endometriosis; endometrial tumour; inflammatory disorder;
KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;
KW ovarian tumour; contact dermatitis; placenta disorder; preeclampsia;
KW EMBRY-2; allergy; gene therapy; gene; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 64..1692
FT CDS /*tag= a
FT /product= "EMBRY-2 protein"

FT sig_peptide 64..123
FT /*tag= b
FT mat_peptide 124..1689
FT /*tag= c
FT /product= "Mature EMBRY-2 protein"
XX
PN WO200248362-A2.
XX
XX
PD 20-JUN-2002.
XX
PF 14-NOV-2001; 2001WO-US043956.
XX
PR 15-NOV-2000; 2000US-0249407P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
XX Ramkumar J, Arvizu C;
XX
XX WPI; 2002-537629/57.
XX P-PSDB; AAE25054.
XX
XX New polypeptides of human embryogenesis associated proteins for screening
XX modulators useful for treating or preventing disorders e.g.
XX endometriosis, infertility, allergy, preeclampsia.
XX
XX Claim 59; Page 96-97; 97pp; English.
XX
XX The invention relates to human embryogenesis associated proteins (EMBRY)
XX and nucleic acid molecules encoding such proteins. EMBRY sequences are
XX useful for screening modulators useful for treating or preventing
XX disorders associated with abnormal expression of EMBRY. The disorders
XX treated include reproductive disorders such as infertility,
XX endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
XX disorder such as acquired immune deficiency syndrome (AIDS), allergies,
XX contact dermatitis; disorders of the placenta such as preeclampsia,
XX abruptio placentae etc. Sequences of the invention are also useful for
XX analysing a proteome of a tissue or a cell type. EMBRY proteins are
XX useful as immunogens for preparing antibodies. Polynucleotides of the
XX invention are useful for creating knockin humanised animals or transgenic
XX animals to model human diseases. They are also used in gene therapy. The
XX present sequence is human EMBRY-2 cDNA
XX
SQ Sequence 2074 BP; 583 A; 567 C; 392 G; 531 T; 0 U; 1 Other;
Query Match 93.8%; Score 1389.8; DB 6; Length 2074;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1424; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
Qy 1 ATGGCCCTCCCTTATCATACTTTCTTCTTACTGTTCTTACCCCTTTGCTCTCACT 60
Db 64 ATGGCCCTCCCTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 123
Qy 61 GCACCCCTCCATGCTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 124 GCACCTCTCCATGCTGTAGGACAGTAGTCCCTTACCAAGCGTTTCTATGGAGA 183
Qy 121 ACGCCGCTTCTGGAAATATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTC 180
Db 184 ATGGCGCTCCCGACAGACATATTGATGCCCATCGTATAGAGTTTATCTAAGGGAAC 243
Qy 181 ACCTTCATGCCCCACACCATCATATGCCCGCACTGTAATCTGCACTCTTTGCTG 240
Db 244 GCCTTTCACGCCCCACACCATATGCCCGCACTGTAATCTGCACTCTTTGCTG 303
Qy 241 CATGCAATACCTATTATGGACAGGAAATGATTAATCTTCTTCTTCTTCTTCT 300
Db 304 CATGCAATACCTATTATGGACAGGAAATGATTAATCTTCTTCTTCTTCTTCT 363
Qy 301 GGAGCCACTGTCTGTTGGAATTTACTTACCCATACAGTATGCTGATGGGGTGAATT 360
Db 364 GGAGCCACTATCTGTTGGAATTTACTTACCCATACCGGTATGCTGATGGGGTGAATT 423
Qy 361 CAAAGTCAGCAAGAGAGAAAACAAGTAAAGGAAGCAATCTCCCACTGACCGGGGACAT 420

Db 424 CAAGATCAGGCAAGAGAAAACATGTAAAGAGTAGTAATCTCCCACTCACCCTGGGTACAT 483
Qy 421 AGACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAGCCCTCCGTACC 480
Db 484 AGCACCCTAGCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAGCCCTCCGTACC 543
Qy 481 CATACTCGCTGTGAGCCTATTATTAACACACCTCAGTGGGTCCATGAGGTCTCGGCC 540
Db 544 CATACTCAGCTGTGAGCCTATTATTAACACACCTCAGTGGGTCCATGAGGTCTCGGCC 603
Qy 541 CAAAACCCCTACTAATCTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
Db 604 CAAAACCCCTACTAATCTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 663
Qy 601 CTGTGCTCTGAAACAATGGAACAACCTTCAGCAGAGAAATAACACCACTTCCTGTTTAGTA 660
Db 664 CTGTGCTCTGAAACAATGGAACAACCTTCAGCAGAGAAATAACACCACTTCCTGTTTAGTA 723
Qy 661 GGACCTCTGCTTCCCAATCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATAATTT 720
Db 724 GSACCTCTGCTTCCCAATCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATAATTT 783
Qy 721 AGCAATCTATAGACACAACAGCTCCCAATGCACTCAGTGGGTAAACACCTCCACACGA 780
Db 784 AGCAATCTATAGACACAACAGCTCCCAATGCACTCAGTGGGTAAACACCTCCACACGA 843
Qy 781 ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 840
Db 844 ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 903
Qy 841 AATGGCTCTTCAGAACTATGTGCTCTCTCTATTTAGTGGCCCTATGACCATCTAC 900
Db 904 AATGGCTCTTCAGAACTATGTGCTCTCTCTATTTAGTGGCCCTATGACCATCTAC 963
Qy 901 ACTGAACAAGATTTATCAATCATGTCTGCTACCTTAAGCCCAACACAAAGAGTACCCATT 960
Db 964 ACTGAACAAGATTTATCAATCATGTCTGCTACCTTAAGCCCAACACAAAGAGTACCCATT 1023
Qy 961 CTTCTTTTGTATCAGAGCAGGAGTCTAGGAGAGCTAGTACTGGCATTTGGCAGTATC 1020
Db 1024 CTTCTTTTGTATCAGAGCAGGAGTCTAGGAGAGCTAGTACTGGCATTTGGCAGTATC 1083
Qy 1021 ACACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGTGACATGGAAACAG 1080
Db 1084 ACACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGTGACATGGAAACAG 1143
Qy 1081 GTCACTGACTCCCTGGTCACTTGCAGAGATCAACTTAACCTCCCTAGCAGAGTAGTCCTT 1140
Db 1144 GTCACTGACTCCCTGGTCACTTGCAGAGATCAACTTAACCTCCCTAGCAGAGTAGTCCTT 1203
Qy 1141 CAAATCGAAGAGCTTTAGACTTCTAACCGCCAAAAGAGGGGAACTGTGTTATTTTAA 1200
Db 1204 CAAATCGAAGAGCTTTAGACTTCTAACCGCCAAAAGAGGGGAACTGTGTTATTTTAA 1263
Qy 1201 GGAGAAGAACGCTGTTATTTATTTAATCAATCCAGAAATGCTCACTGAGAAGTTAAGAA 1260
Db 1264 GGAGAAGAACGCTGTTATTTATTTAATCAATCCAGAAATGCTCACTGAGAAGTTAAGAA 1323
Qy 1261 ATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCGCAACGCTGGGGCTC 1320
Db 1324 ATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCGCAACGCTGGGGCTC 1383
Qy 1321 CTCAGCCAAATCGATGGCCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 1384 CTCAGCCAAATCGATGGCCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG 1443
Qy 1381 TTACTCTCTTTTGAACCCCTGTATCTTTAACTCTCTGTTAGTTGTTCTCTCCAGAAAT 1440
Db 1444 TTACTCTCTTTTGAACCCCTGTATCTTTAACTCTCTGTTAGTTGTTCTCTCCAGAAAT 1503
Qy 1441 GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481

Db 1504 GAAGCTGTAAAGCTACAAATGGTCTTTCAAAATGGAGCCCA 1544

RESULT 8
AAD24195
ID AAD24195 standard; cDNA; 2930 BP.
XX
AC AAD24195;
XX 07-MAY-2002 (first entry)
XX Human syncytin cDNA.
DE Human syncytin; preclampsia; gestational trophoblast disorder;
XX Human syncytin; hydatiform mole; placental site tumour; abortion;
KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;
KW envelope gene; human endogenous defective retrovirus; HERV-W; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 930..2546
FT CDS /*tag= a
FT /product= "Syncytin"
XX
XX WO200204678-A2.
XX
XX 17-JAN-2002.
XX
XX 09-JUL-2001; 2001WO-US021719.
XX
XX 07-JUL-2000; 2000US-0216657P.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Keith JC, Mccoy JM, Mi S;
XX
XX WPI; 2002-171727/22.
XX P-PSDB; AAE14540.
XX
XX Identifying a compound for treating a subject with or at risk of
XX developing preclampsia, comprises determining whether the expression or
XX activity of syncytin in the cell is modulated in the presence of a test
XX compound.
XX
XX Disclosure; Page 39-42; 43pp; English.
XX
XX The invention relates to identifying compounds which are modulators of
XX syncytin expression. The syncytin modulators are useful in diagnosis and
XX treatment of preclampsia and gestational trophoblast disorders (e.g.
XX choriocarcinoma, hydatiform mole, placental site tumour and missed/
XX incomplete abortion). Syncytin is a human gene derived from the envelope
XX gene of human endogenous defective retrovirus, HERV-W. The present
XX invention is based partly on the discovery that syncytin expression is
XX dramatically reduced in preclampsia, and is also mis-localised to the
XX apical syncytiotrophoblast membrane. The present sequence is human
XX syncytin cDNA
XX
SQ Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;

Query Match 90.2%; Score 1335.4; DB 6; Length 2930;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATATCTTTCTTTACTGTTCTTACCCCTTTGGCTCTCACT 60
Db 930 ATGGCCCTCCCTTATCATATCTTTCTTTACTGTTCTTTTACCTCTTTCACTCTCACT 989
Qy 61 GCACCCCTCCATGCTGCTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
Db 990 GCACCCCTCCATGCTGCTGTATGACCAAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 1049
Qy 121 ACGGGGCTTCTCTGGAATATTTGATGCCCCATCATATAGGAGTTTATCTAGGGAACTCC 180

Db 1050 ATGAGCGTCCGGAAATATTGATGCCCATCGTATAGAGTCTTTCTAAGGGAACCCCC 1109
Qy 181 ACCTTCACTGCCACACCATATGCCCCGCAACTGCTATAACTCTGCGCACTCTTTGCAATG 240
Db 1110 ACCTTCACTGCCACACCATATGCCCCGCAACTGCTATAACTCTCTGCCACTCTTTGCAATG 1169
Qy 241 CATGCAATACTCATTTATGGACGAGGAATGATTAATCCCTAGTGTCTCGAGGACCT 300
Db 1170 CATGCAATACTCATTTATGGACGAGGAATGATTAATCCCTAGTGTCTCGAGGACCT 1229
Qy 301 GGAGCCACTGTCTGTGGACTTACTTACCCATACAGTATCTCTGATGGGGGTGGAAT 360
Db 1230 GGAGTCACTGTCTGTGGACTTACTTACCCAACTGGTATCTCTGATGGGGGTGAGTT 1289
Qy 361 CAAAGTTCAGGCAAGAAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGGACAT 420
Db 1290 CAAAGTTCAGGCAAGAAAAAACAAGTAAAGGAAGTAAATCTCCAACTGACCCGGGTACAT 1349
Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCGTACC 480
Db 1350 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGGTACC 1409
Qy 481 CATACTCGCTGTGAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 1410 CATACTCGCTGTGAGCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCGGC 1469
Qy 541 CAAAAACCTACTAATCTGTGGATGTGCTCCCTCCCTGCACCTTCAGGCCATACATTTCAATC 600
Db 1470 CAAAAACCTACTAATCTGTGGATGTGCTCCCTCCCTGAACTTCAGGCCATATGTTCATATC 1529
Qy 601 CCGTGTTCCTGCAATGGAACAACTTCAGACAGAAATAAACAACCACTTCGGTTTTAGTA 660
Db 1530 CCGTGTACCTGAAACAATGGAACAACTTCAGACAGAAATAAACAACCACTTCGGTTTTAGTA 1589
Qy 661 GGACCTCTGTGTTTCCAACTCTGAAATAAACCCATACCTCAAACTCACTCTGTGTAAATTT 720
Db 1590 GGACCTCTGTGTTTCCAACTCTGAAATAAACCCATACCTCAAACTCACTCTGTGTAAATTT 1649
Qy 721 AGCAATACTATAGACAACAACAGCTCCCAATGATCAGGTGGGTAAACACTCCACACGA 780
Db 1650 AGCAATACTATAGACAACAACAGCTCCCAATGATCAGGTGGGTAACTCCTCCCAACAA 1709
Qy 781 ATAGTCTGCTACCTCCAGGAATATTTTGTCTGTGTACCTCAGCCTATCATTTGTTG 840
Db 1710 ATAGTCTGCTACCTCCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTTG 1769
Qy 841 AATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTATGTGCCCCCTATGACCATCTAC 900
Db 1770 AATGGCTCTTCAGAACTATGTGCTTCTCTCATCTTATGTGCCCCCTATGACCATCTAC 1829
Qy 901 ACTGAAACAGATTTATACATATGTGCTACCTAAGCCCCCAACAAGAGGTACCCATT 960
Db 1830 ACTGAAACAGATTTATACATATGTGCTACCTAAGCCCCCAACAAGAGGTACCCATT 1889
Qy 961 CTTTCTTTTGTATCAGACGAGGTGCTAGGCAGACTAGGTACTGGCATTTGGCATATC 1020
Db 1890 CTTTCTTTTGTATCAGACGAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGT 1949
Qy 1021 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAAATGGTGCATGGAACAG 1080
Db 1950 ACAACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAAATGGGACATGGAACG 2009
Qy 1081 GTCACTGATCCCTGTGCTGCAAGATCACTTAATCTCCCTAGCAGCAGTAGTCCCT 1140
Db 2010 GTGCGCGACTCCCTGTGCTGCAAGATCACTTAATCTCCCTAGCAGCAGTAGTCCCT 2069
Qy 1141 CAAAAATCGAAGAGCTTTAGACTTGTAAACCGCCAAAGAGGGGAACTCTTTATTTTAA 1200
Db 2070 CAAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAAGAGGGGAACTCTTTATTTTAA 2129
Qy 1201 GGAGAGAAACGCTGTTATTTATTTAATCAATCAGAAATTTGCTACTGAGAAAGTTAAGAA 1260
Db 2130 GGGGAAAGATGCTGTTATTTATTTAATCAATCAGAAATCGTCACTGAGAAAGTTAAGAA 2189

Qy 1261 ATTGAGATCGAATACAAATGTAGACGAGGCTTCAAAACACCGAACGCTGGGGCCTC 1320
Db 2190 ATTGAGATCGAATACAAACGTAGACGAGGCTTCAAAACACCTGGACCTGGGGCCTC 2249
Qy 1321 CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 2250 CTCAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 2309
Qy 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTAAAGTTTGTCTTCCAGAAAT 1440
Db 2310 CTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTAACTCTTGTAACTTTGTCTTCCAGAAATC 2369
Qy 1441 GAAGCTGTAAAGCTACAGATGCTCTTCAAAATGGAACCCCA 1481
Db 2370 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCCAA 2410

RESULT 9
AAX77526

ID AAX77526 standard; cDNA; 2946 BP.

XX AAX77526;

XX 10-AUG-1999 (first entry)

XX Human secreted protein AJ172_2 cDNA.

XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
cell proliferation; cell differentiation; suppressor; tumour inhibition;
haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.

XX Homo sapiens.

XX WO9926972-A1.

XX 03-JUN-1999.

XX 17-NOV-1998; 98US-05024614.

XX 21-NOV-1997; 97US-00976110.

XX 18-MAY-1998; 98US-00080478.

XX 20-OCT-1998; 98US-00175928.

XX (GEMY) GENETICS INST INC.

XX Jacobs K, McCooy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Merberg D, Treacy M;

XX WPI; 1999-357813/30.

XX P-PSDB; AAY08622.

XX New polynucleotides encoding secreted proteins.

XX Claim 13a; Page 100-101; 142pp; English.

XX This invention describes novel human secreted proteins encoded by
polynucleotides isolated from human adult testes, adult brain, adult
blood or adult placenta, or murine adult bone marrow or thymus cDNA
libraries. The products of the invention are predicted to have biological
activities which would make them suitable for treating, preventing or
ameliorating medical conditions in humans and animals, although no
supporting data is given. Suggested activities include nutritional
activity, cytokine and cell proliferation/differentiation activity,
immune stimulating (e.g. as vaccines) or suppressing activity,
haematopoiesis regulating activity, tissue growth activity,
activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, receptor/ligand activity, anti-inflammatory
activity, cadherin/tumour invasion suppressor activity, and tumour
inhibition activity. The polynucleotides are also stated to be useful for

Db	1888	CTTCTCTTTTGTATAGGACGACGAGTGCTAGGTGCATCTAGGTACTGGCAATTGGCGGTATC	194
Qy	1021	ACAACTCTTACTCAGTTCTTACTACAAACTATCTCRAAGAAATAAATGCTGACATGGGAACAG	1080
Db	1948	ACAACTCTTACTCAGTTCTTACTACAAACTATCTCRAAGAAATAAATGCTGACATGGGAACAG	2007
Qy	1081	GTCACTGACTCCCTGGTGCACCTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT	1140
Db	2008	GTCCGCGACTCCCTGGTGCACCTTTGCAAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT	2067
Qy	1141	CAAATCGAAGAGCTTTAGACTTTGCTTAACCGCCAAAGAGGGGGAACCTGTTTATTTT	1200
Db	2068	CAAATCGAAGAGCTTTAGACTTTGCTTAACCGCTGAAAGAGGGGGAACCTGTTTATTTT	2127
Qy	1201	GGGAAGAAGCGTGTATTATGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAAGAA	1260
Db	2128	GGGAAGAAGCGTGTATTATGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAAGAA	2187
Qy	1261	ATTCGAGATCGAATACAATGTAGACGACGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC	1320
Db	2188	ATTCGAGATCGAATACAATGTAGACGACGAGGAGCTTCAAAACACCGAACGCTGGGGCCTC	2247
Qy	1321	CTCAGCAATGGATGCCCTGGGTCTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
Db	2248	CTCAGCAATGGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	2307
Qy	1381	TTACTCCTCTTTGGACCGCTGTATCTTTAACTCCTTGTGTTAAGTTTGTCTCTCCAGAATT	1440
Db	2308	TTACTCCTCTTTGGACCGCTGTATCTTTAACTCCTTGTGTTAAGTTTGTCTCTCCAGAATT	2367
Qy	1441	GAAGCTGTAAGCTACAGATGGTCTTACAATGGAAACCCCA	1481
Db	2368	GAAGCTGTAAGCTACAGATGGTCTTACAATGGAAACCCCA	2408
RESULT 10			
AAZ59468,			
ID	AAZ59468 standard; cDNA; 2946 BP.		
XX	AAZ59468;		
XX	11-APR-2000 (first entry)		
XX	Human secreted protein AJ172.2 polynucleotide sequence.		
DE	Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;		
XX	placental pathology; metastasis inhibition; nutritional activity;		
KW	immune stimulator; haematopoiesis regulator; tissue growth;		
KW	tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;		
KW	gene therapy; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	WO9960020-A1.		
PN	25-NOV-1999.		
XX	17-MAY-1999; 99WO-US010915.		
PD	18-MAY-1998; 98US-00080478.		
XX	20-OCT-1998; 98US-00175928.		
XX	(GEM) GENETICS INST INC.		
XX	Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;		
PI	Meiberg D, Mi S, Treacy M;		
FI	WPI; 2000-116311/10.		
XX	P-PSDB; AAY67313.		
DR	New polynucleotides encoding secreted cDNA libraries, used to develop		
XX	products for the diagnosis and treatment of neoplastic disease.		
PT			

KW hypersensitivity; regeneration; neural cell proliferation; fertility;
 KW tumour; chemokine; human; secreted protein.
 XX Homo sapiens.

OS US2002193567-A1.

PN 19-DEC-2002.

PD 02-APR-2002; 2002US-00114893.

XX 11-AUG-1995; 95US-00514014.

XX 05-APR-1996; 96US-00628364.

XX 19-APR-1996; 96US-00635311.

XX 07-JUN-1996; 96US-00659224.

XX 17-JUN-1996; 96US-00664596.

XX 09-JUL-1996; 96US-00677231.

XX 26-JUL-1996; 96US-00686878.

XX 23-AUG-1996; 96US-00701819.

XX 27-SEP-1996; 96US-00721488.

XX 27-SEP-1996; 96US-00721798.

XX 27-SEP-1996; 96US-00721923.

XX 27-SEP-1996; 96US-00721926.

XX 25-OCT-1996; 96US-00738367.

XX 30-OCT-1996; 96US-00739775.

XX 13-JAN-1997; 97US-00783395.

XX 10-APR-1997; 97US-00833823.

XX 02-JUN-1997; 97US-00867677.

XX 05-SEP-1997; 97US-00924838.

XX 06-OCT-1999; 99US-00413232.

PA (GEMY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;

PI Kelleher K;

XX WPI; 2003-657236/62.

DR P-PSDB; ADC38777.

XX Proteins A23021 encoded by clone A23021 from human adult colon, and

PT BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,

PT useful for treating e.g. multiple sclerosis and rheumatoid arthritis.

XX Disclosure; SEQ ID NO 134; 412pp; English.

XX The invention relates to a protein comprising fully defined A2302 1

CC protein or BD127 1 6 protein. The polynucleotides are useful for

CC expressing recombinant proteins for analysis and are also useful as

CC chromosome markers or tags to identify chromosomes or to map related gene

CC positions. The proteins are useful as amino acid supplement, carbon

CC source, nitrogen source and carbohydrate source. The proteins are useful

CC for treating various immune deficiencies and disorders (e.g. severe

CC combined immunodeficiency (SCID), autoimmune disorders (e.g. multiple

CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic

CC reactions (e.g. asthma, myeloid or lymphoid cell deficiencies,

CC osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.

CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),

CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory

CC response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease),

CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing

CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or

CC nerve growth or regeneration, for proliferating neural cells and for

CC regenerating nerve and brain tissue, for inducing fertility and for

CC inhibiting tumour growth. Proteins are also useful as chemokine for

CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also

CC useful as inhibitors of receptor/ligand interactions. The present

CC sequence represents cDNA encoding a human secreted protein.

XX Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

XX Query Match 90.2%; Score 1335.4; DB 10; Length 2946;

XX Best Local Similarity 93.9%; Pred. No. 0;

		Matches 1390;	Conservative	0;	Mismatches	91;	Indels	0;	Gaps	0;
QY	1	ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTTACCCCTTTCGCTCACT	60							
DB	928	ATGGCCCTCCCTTATCATATTTTCTCTTTACTGTTTACCCCTTTCACCTCACT	987							
QY	61	GCACCCCTCCATGCTGCTGTACACACAGTAGCTCCCTTACCAAGAGTTCTATGAAGA	120							
DB	988	GCACCCCTCCATGCGCTGTATGACACAGTAGCTCCCTTACCAAGAGTTCTATGGAGA	1047							
QY	121	ACGCGCTTCTGGAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC	180							
DB	1048	ATGACGCTCCCGAATATTTGATGCCCATCTGATAGAGTCTTTCTAAGGGAACCTCC	1107							
QY	181	ACCTTCACTGCCACACCCCATATGCCCCCACTGTATAACTCTCCACTCTTTGCATG	240							
DB	1108	ACCTTCACTGCCACACCCCATATGCCCCCACTGTATAACTCTCCACTCTTTGCATG	1167							
QY	241	CATGCAATACTCATTTATTTGGACAGGGAATGATTAATCCTAGTTGCTCGGAGGACTT	300							
DB	1168	CATGCAATACTCATTTATTTGGACAGGGAATGATTAATCCTAGTTGCTCGGAGGACTT	1227							
QY	301	GGAGCCACTGCTGTTGGACTTACTTCAACCATACCAGTATGTCGATGGGGGTGGAATT	360							
DB	1228	GGAGTCACTGCTGTTGGACTTACTTCAACCATACCAGTATGTCGATGGGGGTGGAATT	1287							
QY	361	CAAGGTCAAGCAAGAGAAAAAAGTAAAGAAAGCAATCTCCCAACTGACCCCGGACAT	420							
DB	1288	CAGATCAGGCAAGAGAAAAACATGTAAAGAAAGTAAATCTCCCAACTGACCCCGGACAT	1347							
QY	421	AGCACCCCTAGCCCCCTACAAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCCGTAAC	480							
DB	1348	GGCACCTCTAGCCCCCTACAAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCGTAAC	1407							
QY	481	CATACCTGCGCTGGTGGCCCTTATTAACACCCCTCACTCGGCTCCATGAGGTCTCAGCC	540							
DB	1408	CATACCTGCGCTGGTAAAGCCCTATTTAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC	1467							
QY	541	CAAAACCTACTAACTGTTGGATGTGCTCCCTCCGACTTCAGGCGCATACATTTCAATC	600							
DB	1468	CAAAACCTACTAACTGTTGGATGTGCTCCCTCCGACTTCAGGCGCATATGTTTCAATC	1527							
QY	601	CTGTTCTCTGAAACAATGGAAACAACCTTCAGCACAGAAATAAACACCTTCGTTTATGTA	660							
DB	1528	CTGTACTCTGAAACAATGGAAACAACCTTCAGCACAGAAATAAACACCTTCGTTTATGTA	1587							
QY	661	GGACCTCTGTTTCCCAATCTGGAATAAACCCATACCTCAAACTCACTGTGTAATAATTT	720							
DB	1588	GGACCTCTGTTTCCCAATCTGGAATAAACCCATACCTCAAACTCACTGTGTAATAATTT	1647							
QY	721	AGCAATACTATAGACACAACACAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGA	780							
DB	1648	AGCAATACTATAGACACAACACAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGA	1707							
QY	781	ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCTCATCTATGTTTG	840							
DB	1708	ATAGTCTGCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCTCATCTATGTTTG	1767							
QY	841	AATGGCTCTTCAGAACTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC	900							
DB	1768	AATGGCTCTTCAGAACTCATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC	1827							
QY	901	ACTGAAACAAGATTTATACAATCATGCTACCTAAAGCCCCCAACAACAAGAGTACCCATT	960							
DB	1828	ACTGAAACAAGATTTATACAATCATGCTACCTAAAGCCCCCAACAACAAGAGTACCCATT	1887							
QY	961	CTTCTCTTTTGTATCAGACAGGAGTCTAGGAGAGATAGGTACTGGCATTTGGCAGTATC	1020							
DB	1888	CTTCTCTTTTGTATAGGAGCAGGAGTCTAGGAGTCTAGGAGTCTAGGAGTCTAGGAGTATC	1947							
QY	1021	ACAACCTCTACTCAGTTCTTACTCAAACTATCTCAAGAAATAAATGTCGATCAATGAAACAG	1080							
DB	1948	ACAACCTCTACTCAGTTCTTACTCAAACTATCTCAAGAAATAAATGTCGATCAATGAAACAG	2007							

1081 GTCACTAGTCTCCTGTCACCTTGCAGATCAACTTAACCTCCTAGCAGCAGTAGTCCTT 1140
 2008 GTGCGCGACTCCTGTCACCTTGCAGATCAACTTAACCTCCTAGCAGCAGTAGTCCTT 2067
 1141 CAAATCGAAGAGCTTTAGACTTGTACCGCAAAAGAGGGGAACTGTTTATTTTA 1200
 2069 CAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGGGGAACTGTTTATTTTA 2127
 1201 GGAGAGAAACGCTGTTTATTTATGTTAATCAATCCAGAAATGTCACGAGAAAGTTAAAGAA 1260
 2128 GGGGAAGATGCTGTTTATTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 2187
 1261 ATTGAGATCGAATCAATGTTAGAGCAGAGAGCTTCAAAACACCGAAGCTGGGGCTC 1320
 2188 ATTGAGATCGAATCAACGTAGAGCAGAGAGCTTCGAAACACTGGACCCCTGGGGCTC 2247
 1321 CTCAGCAATGGATGCGCTGGGTCTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATG 1380
 2248 CTCAGCAATGGATGCGCTGGATCTCCCTCTTCTAGGACCTCTAGCAGCTAATATG 2307
 1381 TTACTCCTCTTTGGACCTGTATCTTTAACTCCTCTGTTTAAAGTTGTCTCTCCAGAAAT 1440
 2308 CTACTCCTCTTTGGACCTGTATCTTTAACTCCTCTGTTTAACTTGTCTCTCCAGAAATC 2367
 1441 GAAGCTGTAAGCTACAGATGCTTCAAAATGGAAACCCCA 1481
 2368 GAAGCTGTAAGCTACAAATGGAGCCCAAGATGCAGTCCAA 2408

RESULT 12
 AAH20070
 ID AAH20070 standard; DNA; 1617 BP.
 AC
 XX
 XX
 DT
 XX
 XX
 DE
 XX
 KW Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein; envelope protein; multiple sclerosis-related superantigen; vaccine; surface antigen; transmembrane; multiple sclerosis; neuroprotective; antisense-therapy; autoimmune disorder; ds.
 XX
 OS Human endogenous retrovirus.
 XX
 FH Key
 FT CDS
 FT
 FT
 FT
 FT
 XX WO200131021-A1.
 XX
 XX
 PD
 PD
 XX
 PF 30-OCT-2000; 2000WO-EP010659.
 XX
 XX
 PR 28-OCT-1999; 99EP-00402690.
 XX
 XX
 PA (UYGE-) UNIV GENEVE.
 XX
 XX
 PI Conrad B, Mach B;
 XX
 XX WPI; 2001-316336/33.
 DR P-PSDB; AAB75138.
 XX
 XX New human retrovirus HERV-W ENV proteins/peptides having superantigen activity useful for diagnosing and treating multiple sclerosis.
 PT
 PT
 XX
 XX Claim 13; Fig 10; 94pp; English.
 PS
 XX
 XX On the basis of the PBS t-RNA motif used for the classification of human

CC endogenous retrovirus (HERVs) the full length endogenous provirus which
 CC was located on the long arm of human chromosome 7 (7q21-22) has been
 CC designated HERV-W. The present invention describes proteins or peptides
 CC of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
 CC have neuroprotective activity, and can be used in: vaccines; antisense-
 CC therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are
 CC useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
 CC disorders. (I) are also useful for identifying substances (and optionally
 CC recovering) capable of binding to a retroviral superantigen associated
 CC with MS, substances capable of blocking SAG activity and substances
 CC capable of blocking transcription or translation of HERV-W retroviral
 CC superantigen. A protein or peptide derived from (I), modified to be
 CC devoid of SAG activity and being capable of generating an immune response
 CC against HERV-W retroviral SAG is useful in therapy. Nucleic acid
 CC molecules encoding (I) are useful as vaccines against MS. Substances
 CC capable of blocking SAG activity, capable of binding to a retroviral
 CC superantigen associated with MS, or capable of blocking transcription or
 CC translation of HERV-W retroviral superantigen for use in treating or
 CC preventing MS, obtained using (I) are useful for the treatment and
 CC prevention of MS. (I) and nucleic acids encoding them are useful for
 CC diagnosing autoimmune disease. The present sequence encodes the
 CC specifically claimed envelope protein of HERV-W designated G
 XX
 SQ Sequence 1617 BP; 442 A; 452 C; 296 G; 427 T; 0 U; 0 Other;
 Query Match 90.1%; Score 1333.8; DB 5; Length 1617;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATCATATCTTTCTTTACTGTCTCTTACCCCTTTGCTCTCACT 60
 DB 1 ATGGCCCTCCCTTATCATATCTTTCTTTACTGTCTTTTACCCCTTTTCACTCTCACT 60
 QY 61 GCACCCCTCATGCTGTGTACACAGTAGTCCCTTACCAGAGTCTTATGAGA 120
 DB 61 GCACCCCTCATGCTGTGTACACAGTAGTCCCTTACCAGAGTCTTATGAGA 120
 QY 121 ACGGGCTTCTGGAAATATTGATGCCCATCATATAGAGTCTTATTAAGGAACTCC 180
 DB 121 ATGAGAGCTCCGGAATATTGATGCCCATCATATAGAGTCTTATTAAGGAACTCC 180
 QY 181 ACCTTCACTGCCCAACCCATATGCCCCGCAACTGCTATACTTGCCACTCTTTCATG 240
 DB 181 ACCTTCACTGCCCAACCCATATGCCCCGCAACTGCTATACTTGCCACTCTTTCATG 240
 QY 241 CATGCAATATCTATTATGACAGGAAATGATTAATCTAGTGTCTCTGAGGACTT 300
 DB 241 CATGCAATATCTATTATGACAGGAAATGATTAATCTAGTGTCTCTGAGGACTT 300
 QY 301 GGAGCCACTGTCTGTGGACTTACTTCCACCATACAGTATGCTCTGATGGGGTGAAT 360
 DB 301 GGAGCCACTGTCTGTGGACTTACTTCCACCATACAGTATGCTCTGATGGGGTGAAT 360
 QY 361 CAAGGTCAGGCAAGAGAAACAACTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
 DB 361 CAAGGTCAGGCAAGAGAAACAACTAAAGGAAGCAATCTCCAACTGACCCGGGACAT 420
 QY 421 AGCACCCTTAGCCCTTACAGAGGACTAGTCTCTCAAACTACATGAAACCTCCGTACC 480
 DB 421 AGCACCCTTAGCCCTTACAGAGGACTAGTCTCTCAAACTACATGAAACCTCCGTACC 480
 QY 481 CATACTCGCTGTGAAGCTATTATTAACACCCCTCACTGGGCTCCATGAGGTCTCGGC 540
 DB 481 CATACTCGCTGTGAAGCTATTATTAACACCCCTCACTGGGCTCCATGAGGTCTCGGC 540
 QY 541 CAAACCCCTTAACTGTTGGATATGCTCCCTGAACTTTCAGGCGCATATGTTTCAATC 600
 DB 541 CAAACCCCTTAACTGTTGGATATGCTCCCTGAACTTTCAGGCGCATATGTTTCAATC 600
 QY 601 CCTGTTCTGCAATGGAACAACTTACAGCAGAGAAATAAACACCACTTCCGTTTTAGTA 660
 DB 601 CCTGTTCTGCAATGGAACAACTTACAGCAGAGAAATAAACACCACTTCCGTTTTAGTA 660

Db 871 CATACTCGCTGGTAGCCTATTTAATACCACTCTCACTGGGCTCATGAGGTCTCGGC 930
Qy 541 CAAAACCCCTACTAACTGTTGGATGCGCTCCCTGCGACTTCAGGCCATACATTTCAATC 600
Db 931 CAAAACCCCTACTAACTGTTGGATGCGCTCCCTGCGACTTCAGGCCATATGTTTCAATC 990
Qy 601 CTGTGCTCTGAACATGGAACAACTTCAGCAGAGAAATAAACACCACTTCCTGTTTAGTA 660
Db 991 CTTGTACCTGAAACAACTGGAACAACTTCAGCAGAGAAATAAACACCACTTCCTGTTTAGTA 1050
Qy 661 GGACCTCTGTTTCCAACTCTGAAATAAACCCATACCTCAAACTCTCACCTGTGTAAATTT 720
Db 1051 GGACCTCTGTTTCCAACTCTGAAATAAACCCATACCTCAAACTCTCACCTGTGTAAATTT 1110
Qy 721 AGCAATATATAGACAAACACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
Db 1111 AGCAATATATAGACAAACACCACTCCCAATGATCAGGTGGGTAAACCTCCACACAA 1170
Qy 781 ATAGTCTGCTACCTCAGGAATATTTTGTGCTGGTACCTCAGCCTATCATTTGTTG 840
Db 1171 ATAGTCTGCTACCTCAGGAATATTTTGTGCTGGTACCTCAGCCTATCATTTGTTG 1230
Qy 841 AATGGCTCTTCAGAACTATGTCCTCTCTCATTTAGTGCCTCTATGACCATCTAC 900
Db 1231 AATGGCTCTTCAGAACTATGTCCTCTCTCATTTAGTGCCTCTATGACCATCTAC 1290
Qy 901 ACTGAACAAGATTTATACAAATCATGTCGTACCTAAGCCCCCAACAAAGAGTACCCATT 960
Db 1291 ACTGAACAAGATTTATACAGTTATGTCTATATCTAAGCCCCCAACAAAGAGTACCCATT 1350
Qy 961 CTTCCCTTTGTTATCAGACGAGAGTCTAGGCAGACTAGGTACTGGCATTCGCGATATC 1020
Db 1351 CTTCCCTTTGTTATAGAGCAGAGTCTAGGTGCACTAGGTACTGGCATTCGCGATATC 1410
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
Db 1411 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1470
Qy 1081 GTCACGTACTCCTGTGCTACCTTGCAGAGTCAACTTAACCTCCCTAGCAGAGTAGTCCTT 1140
Db 1471 GTCGCGACTCCTCGTGCCTTGCAGAGTCAACTTAACCTCCCTAGCAGAGTAGTCCTT 1530
Qy 1141 CAAATCGAGAGCTTTAGACTGCTTAACCGCCAAAGAGGGGACCTGTTTATTTTA 1200
Db 1531 CAAATCGAGAGCTTTAGACTGCTTAACCGCTGAAAGAGGGGAACTGTTTATTTTA 1590
Qy 1201 GGAGAAAGACGCTGTTTATTTATTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA 1260
Db 1591 GGAGAAAGATGCTGTTTATTTATTTATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1650
Qy 1261 ATTGAGATCGAATACAAATGAGAGAGAGGAGCTCAAAAACACCGAAGCTGGGCGCTC 1320
Db 1651 ATTGAGATCGAATACAAATGAGAGAGAGGAGCTTCGAAACACTGGACCTGGGCGCTC 1710
Qy 1321 CTGAGCAATGAGTGCCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATTG 1380
Db 1711 CTGAGCAATGAGTGCCTGGATTCCTCCCTCTTAGGACCTCTAGCAGCTCTAAATATTG 1770
Qy 1381 TTACTCCTCTTTGGACCTGTATCTTTAACTCTCTTTAAAGTTTGTCTCTCCAGAAAT 1440
Db 1771 CTACTCCTCTTTGGACCTGTATCTTTAACTCTCTTTAACTTTGTCTCTCCAGAAATC 1830
Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 1831 GAAGCTGTAAGCTACAAATGGAACCCCAAGATGAGTCCAA 1871

RESULT 14
ABN97927
ID ABN97927 standard; DNA; 2599 BP.
XX
AC ABN97927;

XX 01-AUG-2002 (first entry)
XX Human retroviral sequence HERV 7 env.
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis; ds.
XX Human endogenous retrovirus.
XX WO9967395-A1.
XX 29-DEC-1999.
XX 23-JUN-1999; 99WO-FR001513.
XX 23-JUN-1998; 98FR-00007920.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Alliel PM, Perin J, Rieger F;
XX WPI; 2000-160587/14.
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
XX for diagnosis, treatment and prevention of autoimmune and neurological
XX diseases.
XX Claim 1; Page 128-129; 225pp; French.
XX The present invention relates to new nucleic acid sequences of human
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
XX Regulatory elements associated with HERV-7q may alter expression of other
XX genes (even remote genes) on the same chromosome, inducing immunological
XX and/or neurological changes (which may be pathological or protective/
XX curative). HERV-7q peptides can be used to improve efficiency of the
XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
XX sequences can be used in immunogenic or vaccinating compositions, for
XX protection against autoimmune diseases, particularly multiple sclerosis.
XX The peptides may also be used (by sequence comparison) to detect/identify
XX endogenous retroviruses that are abnormally expressed in cancer,
XX neuropathologies or other autoimmune diseases. The present sequence was
XX used to illustrate the invention
XX Sequence 2599 BP; 744 A; 718 C; 495 G; 642 T; 0 U; 0 Other;
SQ
Query Match 90.1%; Score 1333.8; DB 3; Length 2599;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
Qy 1 ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTCTCTTACCCCTTTCTCTCACT 60
Db 929 ATGGCCCTCCCTTATCATATTTTCTTTACTGTCTTTTACCTCTTTCACTCTCACT 988
Qy 61 GCACCCCTCCATGCTGTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 989 GCACCCCTCCATGCTGTGTATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGGAGA 1048
Qy 121 AGCGGGCTTCGGAAATATTGATGCCCAATCATATAGGAGTTTATCTAAGGGAACCTCC 180
Db 1049 ATGACGCTGCCGAAATATTGATGCCCAATCATATAGGAGTTTCTAAGGGAACCTCC 1108
Qy 181 ACCTTCACTGCCACACCATATGCTCCGCAACTGCTATACTGCCACTCTTTGCAATG 240
Db 1109 ACCTTCACTGCCACACCATATGCTCCGCAACTGCTATACTCTGCCACTCTTTGCAATG 1168
Qy 241 CATGCAAAATCACTATTATGGACAGGGAATAATGATTAATCTTAGTTGCTCGAGGACTT 300
Db 1169 CATGCAAAATCACTATTATGGACAGGGAATAATGATTAATCTTAGTTGCTCGAGGACTT 1228
Qy 301 GGAGCCTGCTGTGTGGACTTACTTCCACCCATACAGTATCTCTGATGGGGTGGAAAT 360
Db 1229 GGAGCCTGCTGTGTGGACTTACTTCCACCCAACTGGTATGTCTGATGGGGTGGAGTT 1288

QY 361 CAAGGTCAGGACGAGAAACAAAGTAAGGAGCAATCTCCCAACTGACCCGGGACAT 420
Db 1289 CAAGATCAGGCAAGAGAAACAAATGTAAGAAAGTAATCTCCCACTCACCAGGATACAT 1348
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
Db 1349 GGCACCTCTAGCCCTCAAAAGGACTAGATCTCTCAAAACTACATGAACCCCTCCGTACC 1408
QY 481 CATATCGGCTGTGAGCCTATTATATACACCCCTCACTCGGCTCATGAGTCTCAGCC 540
Db 1409 CATACTCGCTGTGAGCCTATTATATACACCCCTCACTCGGCTCATGAGTCTCAGCC 1468
QY 541 CAAAACCCCTACTACTGTGGATGTCCTCCCTCGCTCACTTCAGGCAATACATTTCAATC 600
Db 1469 CAAAACCCCTACTACTGTGGATGTCCTCCCTCGCTCACTTCAGGCAATATGTTTCAATC 1528
QY 601 CTGTGTTCTGAACAATGGAACTCTCAGCACAGAAATAAACCACTTCCTGTTTAGTA 660
Db 1529 CTGTGTTCTGAACAATGGAACTCTCAGCACAGAAATAAACCACTTCCTGTTTAGTA 1588
QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAAATTT 720
Db 1589 GGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAAATTT 1648
QY 721 AGCAATACTATAGACACACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
Db 1649 AGCAATACTATAGACACACAGCTCCCAATGATCAGGTGGGTAACTCCTCCACACAA 1708
QY 781 ATAGTCGCTTACCCCTCAGGAATATTTTTGTGTGTGCTACCTCAGCTATCATTTGTTG 840
Db 1709 ATAGTCGCTTACCCCTCAGGAATATTTTTGTGTGTGCTACCTCAGCTATCATTTGTTG 1768
QY 841 AATGGCTCTTCAAGATCTATGTGCTCTCTCACTTATGTCCTCCCTATGACCATCTAC 900
Db 1769 AATGGCTCTTCAAGATCTATGTGCTCTCTCACTTATGTCCTCCCTATGACCATCTAC 1828
QY 901 ACTGAAACAGTTTATCAATCATGTCGTACCTTAAGCCCAACAAAGAGTACCCATT 960
Db 1829 ACTGAAACAGTTTATCAAGTATGTCATATCTAAGCCCAACAAAGAGTACCCATT 1888
QY 961 CTTCCTTTTGTATCAGACGAGGAGTCTAGGCAGACTAGTACTGCGATTGGCAGTATC 1020
Db 1889 CTTCCTTTTGTATCAGACGAGGAGTCTAGGCAGACTAGTACTGCGATTGGCAGTATC 1948
QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTTGACATGGAACAG 1080
Db 1949 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATGTTGGAACATGGAACGG 2008
QY 1081 GTCACTGACTCCCTGGTCACTTCAAGATCACTTAACCTCCCTAGCAGCAGTATGCTCT 1140
Db 2009 GTCCCGACTCCCTGGTCACTTCAAGATCACTTAACCTCCCTAGCAGCAGTATGCTCT 2068
QY 1141 CAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAGAGGGGAACTGTTTATTTTAA 1200
Db 2069 CAAATCGAAGAGCTTTAGACTTCTTAACCGCTGAAGAGGGGAACTGTTTATTTTAA 2128
QY 1201 GGAGAAGAACGCTGTTATGTTTAAATCAATPCCAGATGTCATGAGAAAGTTAAAGAA 1260
Db 2129 GGGGAAGAAATGCTGTTTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 2188
QY 1261 ATTGAGATCGAATACATGATGAGCAGAGAGCTTCAAAACCGGACCGCTGGGGCTC 1320
Db 2189 ATTCGAGATCGAATACACGATGAGCAGAGAGCTTCAAAACCTGGACCTGGGGCTC 2248
QY 1321 CTCAGCAATGATGCTGGTCTTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTTG 1380
Db 2249 CTCAGCAATGATGCTGGTCTTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTTG 2308
QY 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTGTTAGTTGTTCTTCTCCAGAAAT 1440
Db 2309 CTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTGTTAGTTGTTCTTCTCCAGAAAT 2368

QY 1441 GAAGCTGTAAAGCTACAGATGCTTTCACAAATGGAACCCCA 1481
Db 2369 GAAGCTGTAAAGCTACAAATGGAACCCCAAGATGCAATGCCAA 2409
RESULT 15
ADS30988
ID ADS30988 standard; DNA; 8523 BP.
XX AC ADS30988;
XX DT 18-NOV-2004 (first entry)
XX Human genome high complexity repeat found in the HIRA gene #21.
XX Human; ds;
KW histone cell cycle regulation defective, S. cerevisiae homologue A; HIRA;
KW high complexity repeat; in situ hybridisation; Southern blot;
KW chromosome breakpoint; inherited genetic disease; neoplastic disorder;
KW chromosome 22; DiGeorge syndrome; Velo-Cardio-facial syndrome.
XX Homo sapiens.
XX US2003224356-A1.
XX 04-DEC-2003.
XX 14-MAY-2001; 2001US-00854867.
XX 16-MAY-2000; 2000US-00573080.
XX (KNOL/) KNOLL J H M.
XX (ROGA/) ROGAN P K.
XX Knoll JHM, Rogan PK;
XX WPI; 2002-062378/08.
Single copy genomic hybridization probes for detecting specific nucleic acid sequences in sample by in situ hybridization useful for detection of acquired or inherited genetic diseases.
Example 1; SEQ ID NO 21; 30pp; English.
The invention relates to a nucleic acid hybridisation probe comprising a labelled, single copy nucleic acids of at least 50 nucleotides, which will hybridise to a deduced single copy sequence interval in target nucleic acid (TNA) of known sequence. The single copy sequence is deduced by comparing the target nucleic acid (e.g. a disease causing gene) with a collection of high and low complexity repeat sequences as found in the genome of the organism from containing the target nucleic acid. The probe is generated by PCR on the target sequence. The probe is essentially free of blocking nucleic acid sequences which will hybridise to repeat sequences within the genome of which the TNA is a part, and is labelled with a label selected from fluorochrome-responsive labels, fluorochromes, calorimetric chemical, conjugated proteins, antibodies, antigens and their mixtures. The probe is useful in a hybridisation method, where the hybridisation method is from in situ hybridisation, Southern blot, and other methods in which nucleic acid is immobilised, where the method further comprises selecting a single copy nucleic acid which will hybridise to a duplicon or triplicon sequence domain. The probe is useful for determining the existence of previously unknown repeat sequence families in a genome. The method comprises reacting a labelled probe with the genome, causing the probe to hybridise and ascertaining if the probe hybridises to the genome at more than three preferably ten different locations as a determination of new repeat sequence family, where the determining step comprises selecting the single copy sequence from a duplicon or triplicon sequence domain. The probe is useful for determining a chromosome breakpoint and is useful in the fields for cytogenetics and molecular genetics for determining the presence of specific nucleic acid sequences in a sample of eukaryotic origin, e.g. the probes may be used to analyse specific chromosomal locations by in situ hybridisation as a detection of acquired or inherited genetic

CC diseases especially for detection of genetic or neoplastic disorders.
CC Unlike prior art techniques, the probe permits more precise chromosomal
CC breakpoint determinations by in situ hybridisation. The genomic sequence
CC comprising the human HIRA gene (histone cell cycle regulation defective,
CC S. cerevisiae, homolog A) was analysed for single copy sequence
CC intervals for use as probes of the invention. HIRA is located on
CC chromosome 22 as a duplicate, deletions of 1 copy lead to DiGeorge and
CC Velo-Cardio-facial syndromes. The present sequence is a high complexity
CC repeat found within the human genome used to analyse the HIRA gene for
CC repeat regions. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030224356.
XX
XX

Sequence 8523 BP; 2560 A; 2118 C; 1814 G; 2031 T; 0 U; 0 Other;

Query Match	90.1%;	Score 1333.8;	DB 7;	Length 8523;
Best Local Similarity	93.8%;	Pred. No. 0;		
Matches 1389;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;
QY	1	ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTTCTCTTACCCCTTTTGGCTCTCACT	60	
DB	6849	ATGGCCCTCCCTTATCATATTTTCTTTTACTGTTCTTTTACCCCTTTTTCACCTCTCACT	6908	
QY	61	GCACCCCTCCATGCTGCTGTACACACAGTACCTCCCTTACCAGAGTTTCTATGAAGA	120	
DB	6909	GCACCCCTCCATGCGCGCTGTATGACACAGTAGCTCCCTTACCAGAGTTTCTATGGAGA	6968	
QY	121	ACCGCGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACTCC	180	
DB	6969	ATCGAGGTTCCGGAATATTGATGCCCATCGTATAGGAGTTTCTAAGGGAACTCC	7028	
QY	181	ACCTTCACCTGCCACACCCATATGCCCCGGAACCTGCTATTAACCTCGCCTCTTTGCA	240	
DB	7029	ACCTTCACCTGCCACACCCATATGCCCCGGAACCTGCTATTAACCTCGCCTCTTTGCA	7088	
QY	241	CATGCAAACTCATTTATGGACAGGAAATGATTAATCTATGTTGCTCTGAGGACTT	300	
DB	7089	CATGCAAACTCATTTATGGACAGGAAATGATTAATCTATGTTGCTCTGAGGACTT	7148	
QY	301	GGAGCCACTCTCTGTTGGACTTACTTACCCCAACTGGTATGTCATGAGGGTGGAA	360	
DB	7149	GGAGTCACTCTCTGTTGGACTTACTTACCCCAACTGGTATGTCATGAGGGTGGAG	7208	
QY	361	CAAGGTGAGGAG	420	
DB	7209	CAAGATCAGGCAAG	7268	
QY	421	AGCACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAAACTACATGAAACCTCCGTACC	480	
DB	7269	GGCACCTCTAGCCCTACAAAGGACTAGTTTCTCTCAAAACTACATGAAACCTCCGTACC	7328	
QY	481	CATACTCGCTGTTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCTAGCC	540	
DB	7329	CATACTCGCTGTTGAGCTTATTAATACACCCCTCACTCGGCTCCATGAGGTCTAGCC	7388	
QY	541	CAAAACCCCTACTACTGTTGATGTCCTCCCTGCACTTCAGGCCATACATTTCAATC	600	
DB	7389	CAAAACCCCTACTACTGTTGATGTCCTCCCTGCACTTCAGGCCATATGTTTCAATC	7448	
QY	601	CCTGTTCTGAAACAACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	660	
DB	7449	CCTGTTCTGAAACAACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	7508	
QY	661	GGACCTCTTGTTCCTTCAATCTGGAATAACCCATACCTCAAACTCAGCTGTGTAATTT	720	
DB	7509	GGACCTCTTGTTCCTTCAATCTGGAATAACCCATACCTCAAACTCAGCTGTGTAATTT	7568	
QY	721	AGCAATACTATAGCAACACAGCTCCCAATGCAATCAGGTGGGTAAACCTCCACACGA	780	
DB	7569	AGCAATACTATAGCAACACAGCTCCCAATGCAATCAGGTGGGTAACTCTCCACACAA	7628	
QY	781	ATAGTCTGCTCCCTCAGGAAATATTTTGTCTGTGTACCTCAGCTATCATGTTTG	840	

Db	7629	ATAGTCTGCTCCCTCAGGAAATATTTTGTCTGTGTACCTCAGCTATCGTTGTTG	7688	
QY	841	AATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900	
Db	7689	AATGGCTCTTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	7748	
QY	901	ACTGAAACAAGATTATACAATCATGTCTACCTTAAGCCCCACAAACAAAGAGTACCCATT	960	
Db	7749	ACTGAAACAAGATTATACAAGTATGCTATATCTTAAGCCCCACAAACAAAGAGTACCCATT	7808	
QY	961	CTTCTCTTTTGTATTCAGAGCAGAGAGTCTAGGAGAGAGTACTGAGTACTGGCAGTATC	1020	
Db	7809	CTTCTCTTTTGTATTCAGAGCAGAGAGTCTAGGAGAGAGTACTGAGTACTGGCAGTATC	7868	
QY	1021	ACAACCTCTACTAGTTTCTACTACAACTATCTCAAGAAATTAATGTGACATGGAACAG	1080	
Db	7869	ACAACCTCTACTAGTTTCTACTACAACTATCTCAAGAAATTAATGTGGAACATGGAACAG	7928	
QY	1081	GTCACTGACTCCCTGGTCACTTTCGAAAGATCAACTTAACCTCCCTAGCAGAGTACTGCTT	1140	
Db	7929	GTCCGCGACTCCCTGGTCACTTTCGAAAGATCAACTTAACCTCCCTAGCAGAGTACTGCTT	7988	
QY	1141	CAAAATCGAAGAGCTTTAGACTTCTAACCCGCTGAAAGAGGGGGAACCTGTTTATTTT	1200	
Db	7989	CAAAATCGAAGAGCTTTAGACTTCTAACCCGCTGAAAGAGGGGGAACCTGTTTATTTT	8048	
QY	1201	GGAGAGAACCTGTTTATTTATTTAATCAATCCAGAAATTTGCTACTGAGAAAGTTAAAGAA	1260	
Db	8049	GGGGAAGAAATGCTGTTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA	8108	
QY	1261	ATTGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAAACACGAAACGCTGGGGCTC	1320	
Db	8109	ATTGAGATCGAATACAACTAGAGCAGAGAGCTTCAAAAACACGAAACGCTGGGGCTC	8168	
QY	1321	CTCAGCCAAATGATGCTCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380	
Db	8169	CTCAGCCAAATGATGCTCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	8228	
QY	1381	TTACTCTCTTTGGACCTGATCTTTAACTCTCTTTTAAAGTTTGTCTCTTCCAGAAAT	1440	
Db	8229	CTACTCTCTTTGGACCTGATCTTTAACTCTCTTTTAACTCTCTTTTAACTTTGTCTCTTCCAGAAATC	8288	
QY	1441	GAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA	1481	
Db	8289	GAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAAGTCCAA	8329	
RESULT 16				
ABN97929				
ID	ABN97929	standard; DNA; 10499 BP.		
XX	AC	ABN97929;		
XX	XX			
XX	DT	01-AUG-2002 (first entry)		
XX	DE	Human retroviral sequence HERV-7q.		
XX	XX	Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;		
XX	KW	multiple sclerosis; ds.		
XX	OS	Human endogenous retrovirus.		
XX	XX	WO9967395-A1.		
XX	PN			
XX	PD	29-DEC-1999.		
XX	XX			
XX	PF	23-JUN-1999; 99WO-FR001513.		
XX	XX			
XX	PR	23-JUN-1998; 98FR-00007920.		
XX	XX			
XX	PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.		
XX	PI	Alliel PM, Perin J, Rieger F;		

CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
SQ Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;
Query Match 90.1%; Score 1333.8; DB 6; Length 56093;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTTATCATCTCTTTCTTTACTCTCTTACCTCTCTTACCCCTTTCCGCTCTCACT 60
DB 35879 ATGGCCCTCCCTTATCATCTCTTTCTTTACTGTCTTTTACCCTCTTTTACCTCTCACT 35938
QY 61 GCACCCCTCCATCTGCTGTACCAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 35939 GCACCCCTCCATGCGCTGTATGCCAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 35998
QY 121 ACGCGGCTTCTGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAACTCC 180
DB 35999 ATGCAGCTGCCGGAATATTGATGCCCATCATATAGGAGTTTCTTAAAGGAACTCC 36058
QY 181 ACCTTCACTGCGCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 240
DB 36059 ACCTTCACTGCGCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 36118
QY 241 CATGCAATATCTCAATTATTGGACAGGGAATAATGATTAATCTTAGTTCTCTGGAGGACTT 300
DB 36119 CATGCAATATCTCAATTATTGGACAGGGAATAATGATTAATCTTAGTTCTCTGGAGGACTT 36178
QY 301 GGAGCACTGTCTGTGGACCTTATCTTACCATACCACTATGCTGTGATGGGGTGGAAAT 360
DB 36179 GGAGCACTGTCTGTGGACCTTATCTTACCATACCACTATGCTGTGATGGGGTGGAGTT 36238
QY 361 CAAGTCAAGGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGAGCAT 420
DB 36239 CAAGTCAAGGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGAGCAT 36298
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCTCGTACC 480
DB 36299 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCTCGTACC 36358
QY 481 CATACTCGCTGTGGAGCTTATTAATACCACTCTCTGCTCCATGAGGTCTCAGCC 540
DB 36359 CATACTCGCTGTGGAGCTTATTAATACCACTCTCTGCTCCATGAGGTCTCAGCC 36418
QY 541 CAAAACCTTACTAATCTGTGGATGTCCTCCCTCTGCACCTTCCAGGCCATACATTTCAATC 600
DB 36419 CAAAACCTTACTAATCTGTGGATGTCCTCCCTCTGCACCTTCCAGGCCATACATTTCAATC 36478
QY 601 CCTGTTCTGAAACAATGGAACAACCTTACGACAGAAATAAACAACCACTTCCTGTTTAGTA 660
DB 36479 CCTGTTCTGAAACAATGGAACAACCTTACGACAGAAATAAACAACCACTTCCTGTTTAGTA 36538
QY 661 GGACCTCTGTTTCAATCTGGAATAACCACTACCTCAAACTCCTGCTGTGTAATTT 720
DB 36539 GGACCTCTGTTTCAATCTGGAATAACCACTACCTCAAACTCCTGCTGTGTAATTT 36598
QY 721 AGCAATCTATAGACAAACAGCTCCCAATCATCATGAGTGGGTAAACCTCCCAACACA 780
DB 36599 AGCAATCTATAGACAAACAGCTCCCAATCATCATGAGTGGGTAAACCTCCCAACACA 36658
QY 781 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTT 840
DB 36659 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTGGTACCTCAGCCTATCATTTGTTT 36718
QY 841 AATGGCTCTTCAGATCTATGCTTCTCTCTATTTAGTCCCTTATGACCTATGACCTATAC 900
DB 36719 AATGGCTCTTCAGATCTATGCTTCTCTCTATTTAGTCCCTTATGACCTATGACCTATAC 36778

XX WO200194629-A2.
XX 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
XX 18-SEP-2000; 2000US-0233133P.
XX 18-SEP-2000; 2000US-0233161P.
XX 20-SEP-2000; 2000US-0234009P.
XX 20-SEP-2000; 2000US-0234034P.
XX 20-SEP-2000; 2000US-0234052P.
XX 22-SEP-2000; 2000US-0234509P.
XX 22-SEP-2000; 2000US-0234567P.
XX 25-SEP-2000; 2000US-0234923P.
XX 25-SEP-2000; 2000US-0234924P.
XX 25-SEP-2000; 2000US-0235077P.
XX 25-SEP-2000; 2000US-0235082P.
XX 25-SEP-2000; 2000US-0235134P.
XX 25-SEP-2000; 2000US-0235280P.
XX 26-SEP-2000; 2000US-0235637P.
XX 26-SEP-2000; 2000US-0235638P.
XX 27-SEP-2000; 2000US-0235711P.
XX 27-SEP-2000; 2000US-0235720P.
XX 27-SEP-2000; 2000US-0235840P.
XX 27-SEP-2000; 2000US-0235863P.
XX 28-SEP-2000; 2000US-0236028P.
XX 28-SEP-2000; 2000US-0236032P.
XX 28-SEP-2000; 2000US-0236033P.
XX 28-SEP-2000; 2000US-0236034P.
XX 28-SEP-2000; 2000US-0236109P.
XX 28-SEP-2000; 2000US-0236111P.
XX 29-SEP-2000; 2000US-0236842P.
XX 29-SEP-2000; 2000US-0236891P.
XX 02-OCT-2000; 2000US-0237172P.
XX 02-OCT-2000; 2000US-0237173P.
XX 02-OCT-2000; 2000US-0237278P.
XX 02-OCT-2000; 2000US-0237295P.
XX 02-OCT-2000; 2000US-0237316P.
XX 02-OCT-2000; 2000US-0237425P.
XX 03-OCT-2000; 2000US-0237598P.
XX 03-OCT-2000; 2000US-0237604P.
XX 03-OCT-2000; 2000US-0237606P.
XX 03-OCT-2000; 2000US-0237608P.
XX 01-NOV-2000; 2000US-0244867P.
XX 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 81; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result

QY	661	GGACCTCTGTTTCCAACTCGGAATAACCCATACCTCAAACTCACCTGTGTAAAAATTT	720					
Db	1422	GGACCTCTGTTTCCAACTCGGAATAACCCATACCTCAAACTCACCTGTGTAAAAATTT	1481					
QY	721	AGCAATACATATAGACACACAGCTCCCAATGCATCAGGTGGTAAACCTCCACACGA	780					
Db	1482	AGCAATACATATACACACACCAACTCCCAATGCATCAGGTGGTAACTCCTCCACACAA	1541					
QY	781	ATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGTACTCAGCCTATCATTTGTTTG	840					
Db	1542	ATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGTACTCAGCCTATCATTTGTTTG	1601					
QY	841	AATGGCTCTTCAGAATCTATGCTTCCCTCATCTTTAGTGCCCTCATGACCATCTAC	900					
Db	1602	AATGGCTCTTCAGAATCTATGCTTCCCTCATCTTTAGTGCCCTCATGACCATCTAC	1661					
QY	901	ACTGAACAAGATTTATACAAATCATGTCGTACCTAAAGCCCAACAAAGAGTACCCATT	960					
Db	1662	ACTGAACAAGATTTATACAGTTATGTCTATCTAAGCCCGCAACAAAGAGTACCCATT	1721					
QY	961	CTTCTCTTTTGTATCAGACGAGGTGCTAGGACACTAGGTACTGGCAATGGCAGTATC	1020					
Db	1722	CTTCTCTTTTGTATAGGAGCGGAGTGTAGTGCACTAGGTACTGGCAATGGCAGTATC	1781					
QY	1021	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG	1080					
Db	1782	ACAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGGAACGG	1841					
QY	1081	GTCACGTACTCCTGGTCACTTGCAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCCTT	1140					
Db	1842	GTCCGCGACTCCCTGGTCACTTGCAGAGTCAAGTAACTCCCTAGCAGCAGTAGTCCTT	1901					
QY	1141	CAAAATCGAAGCTTTAGACTTGCTTAACCGCAAGAGGGGAACTGTTTATTTTAA	1200					
Db	1902	CAAAATCGAAGCTTTAGACTTGCTTAACCGCTGAAGAGGGGAACTGTTTATTTTAA	1961					
QY	1201	GGAGAAAGACGCTGTTTATTTATGTTAATCAATCCAGAAATGTTCACGTGAGAAAGTTAAAGAA	1260					
Db	1962	GGGAAAGATGCTGTTTATTTATGTTTAAATCAATCCGAATGCTCACTGAGAAAGTTAAAGAA	2021					
QY	1261	ATTGAGATCGAATACAAATGTAGACGAGGAGCTTCAAAACACCGAAGCTGGGGCCTC	1320					
Db	2022	ATTGAGATCGAATACAACTAGACGAGGAGCTTCAAAACACACTGGACCTGGGGCCTC	2081					
QY	1321	CTCAGCAATGAGTGCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380					
Db	2082	CTCAGCAATGAGTGCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTAATATTG	2141					
QY	1381	TTACTCTCTTTGGACCTGTATCTTTTAACTCTTGTAAAGTTTGTCTCTTCCAGAAAT	1440					
Db	2142	CTACTCTCTTTGGACCTGTATCTTTTAACTCTTGTAACTTGTCTCTTCCAGAAATC	2201					
QY	1441	GAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAAACCCCA	1481					
Db	2202	GAAGCTGTAAACTACAAATGGAGCCCAAGATGCACTCCA	2242					
RESULT 19								
AAA59215								
ID	AAA59215 standard; DNA; 7582 BP.							
XX								
AC	AAA59215;							
XX								
DT	07-NOV-2000 (first entry)							
XX								
DE	Human endogenous retrovirus W (HERV-W) sequence.							
XX								
KW	Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;							
KW	gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.							
XX								
OS	Human endogenous retrovirus.							
XX								
FH	Key							
Location/Qualifiers								
LTR	1..120							
FT	/*tag= a							
FT	/note= "R of 5' LTR"							
LTR	121..575							
FT	/*tag= b							
FT	/note= "US of 5' LTR"							
primer_bind	579..596							
FT	/*tag= c							
FT	5581..7194							
FT	/*tag= d							
FT	/note= "ORF1 env538"							
CDS	7039..7194							
FT	/*tag= e							
FT	/note= "ORF2 52 AA"							
CDS	7112..7255							
FT	/*tag= f							
FT	/note= "ORF3 48 AA"							
misc_feature	7244..7254							
FT	/*tag= g							
FT	/note= "polypurine tract"							
LTR	7256..7582							
FT	/*tag= h							
FT	/note= "U3-R of 3' LTR"							
polyA_signal	7563..7569							
FT	/*tag= i							
WO200043521-A2.								
PD	27-JUL-2000.							
XX								
PF	21-JAN-2000; 2000WO-FR000144.							
XX								
PR	21-JAN-1999; 99FR-00000888.							
XX								
PA	(INMR) BIO MERIEUX.							
XX								
PI	Paranhos-Baccala G, Mallet F, Voisset C;							
XX								
DR	WPI; 2000-49929/44.							
XX								
PT	New nucleic acid from human endogenous retrovirus, useful e.g. for							
PT	diagnosis of autoimmune disease and complications of pregnancy, contains							
PT	at least part of the gag gene.							
XX								
PS	Disclosure; Page 49-52; 53pp; French.							
XX								
CC	The present sequence represents an endogenous retrovirus, which is							
CC	associated with an autoimmune disease, and is integrated into the human							
CC	genome. The retrovirus is human endogenous retrovirus W (HERV-W). The							
CC	HERV-W retrovirus is associated with autoimmune disease, failure of							
CC	pregnancy or disorders of pregnancy. HERV-W nucleic acid fragments, or							
CC	proteins derived from it, are useful for diagnosis of autoimmune disease							
CC	(specifically multiple sclerosis) and for monitoring pregnancy. The							
CC	nucleic acid fragments may also be used for in situ labelling of isolated							
CC	chromosomes, while the transcription product can be used to study or							
CC	monitor T cell proliferation in vitro							
XX								
SQ	Sequence 7582 BP; 2156 A; 1876 C; 1538 G; 1796 T; 0 U; 216 Other;							
Query Match 89.4%; Score 1324.2; DB 3; Length 7582;								
Best Local Similarity 92.2%; Pred. No. 0;								
Matches 1365; Conservative 24; Mismatches 92; Indels 0; Gaps 0;								
QY	1	ATGGCCCTCCCTTATCATATCTTTCTCTTACTGTTCTTACCCCTTTCCGCTCTCACT	60					
Db	5581	ATGGSCCTCCCTTATCATATATTTTCTKTAGTGTSTTTTACCTTTTCACTCTCACT	5640					
QY	61	GCACCCCTCCATGCTGCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA	120					
Db	5641	GCACCCCTCCATGCTGCTGTATGACCACTAGTCCCTTACCMAGAGTTTCTATGGAGA	5700					
QY	121	ACGCGGCTCCCTGGAAATATTGATCCCATCATAGGATTTATCTAAGGGAATCTC	180					

5701 ATGCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTSTAAGGAAACCCCC 5760
 181 ACCTTCACCTGCCACACCATATATGCCCGCAACTGCTATTAATCTGCGCACTCTTTGCAATG 240
 5761 ACCTTCACCTGCCACACCATATATGCCCGCAACTGCTATTAATCTGCGCACTCTTTGCAATG 5820
 241 CATGCAAAATACTCATTTATGGACAGGGAATAATGATTAACTCTAGTTGTCTCGAGGACTTT 300
 5821 CATGCAAAATACTCATTTATGGACAGGGAATAATGATTAACTCTAGTTGTCTCGAGGACTTT 5880
 301 GGAGCCACTGTCTGTGGACTTACTTACCCTATACCATGATGATGCTGATGGGGTGAATTT 360
 5881 GGAGTCACTGTCTGTGGACTTACTTACCCTATACCATGATGATGCTGATGGGGTGAATTT 5940
 361 CAAGGTGAGCAAGAGAAAAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
 5941 CAAGATCAGCAAGAGAAAAAAGTAAAGGAAGTAACTCTCCCACTGACCCGGGTACAT 6000
 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
 6001 GGCACTCTAGCCCTACAAAGGACTAGTCTCTCAAACTACATGAAACCCCTCCGTACC 6060
 481 CATACTCGCTGTGAGCTTATTTAATACACACCTCACTCGGCTCCATGAGTCTCAGCC 540
 6061 CATACTCGCTGTGAGCTTATTTAATACACACCTCACTCGGCTCCATGAGTCTCAGCC 6120
 541 CAAAACCTCTAATCTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
 6121 CAAAACCTCTAATCTGTGGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 6180
 601 CCGTGTCTCAACAAATGGAACAATCTCAGCACAGAAATAAACACCACTTCGCTTTTAGTA 660
 6181 CCGTGTCTCAACAAATGGAACAATCTCAGCACAGAAATAAACACCACTTCGCTTTTAGTA 6240
 661 GGACCTCTTGTCTTCCAACTGGAATAACCCATCACTCAAACTCACTCTGTGTAATTT 720
 6241 GGACCTCTTGTCTTCCAACTGGAATAACCCATCACTCAAACTCACTCTGTGTAATTT 6300
 721 AGCAATCTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
 6301 AGCAATCTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 6360
 781 ATAGTCTGCTACCTCAGCAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 6361 ATAGTCTGCTACCTCAGCAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6420
 841 AATGGCTCTTCAGAACTATGTGTCTCTCTCACTTCTAGTGGCCCTATGACCATCTAC 900
 6421 AATGGCTCTTCAGAACTATGTGTCTCTCTCACTTCTAGTGGCCCTATGACCATCTAC 6480
 901 ACTGAAACAAGATTTATCAATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 6481 ACTGAAACAAGATTTATCAATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6540
 961 CTTCTCTTTCTATCAGCAGGAGTGTAGGAGACTAGTACTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 6541 CTTCTCTTTCTATCAGCAGGAGTGTAGGAGACTAGTACTGCTGCTGCTGCTGCTGCTGCTGCT 6600
 1021 ACAACCTCTACTCAGTGTCTACTACAACTATCTCAAGAAATAAATGATGATGAGCAAG 1080
 6601 ACAACCTCTACTCAGTGTCTACTACAACTATCTCAAGAAATAAATGATGATGAGCAAG 6660
 1081 GTCACTGACTCCCTGGTCACTTGCAGAGATCACTTAATCTCCCTAGCAGCAGTGTCTCTT 1140
 6661 GTCCGCGACTCCCTGGTCACTTGCAGAGATCACTTAATCTCCCTAGCAGCAGTGTCTCTT 6720
 1141 CAAATCGAAGCTTTAGACTTCTTACCCGCAAGAGGGGAACTGTTTATTTT 1200
 6721 CAAATCGAAGCTTTAGACTTCTTACCCGCAAGAGGGGAACTGTTTATTTT 6780
 1201 GGAGAAGACCTGTTTATTTATTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
 6781 GGGGAGAAATGCTGTTTATTTAATCAATCCGAAATCGTCACTGAGAAAGTTAAAGAA 6840

1261 ATTCAGATCGAATAAATGTAGACGAGGAGCTTCAAAAACACCGAAGCGTGGGCGCTC 1320
 6841 ATTCAGATCGAATAAATGTAGACGAGGAGCTTCAAAAACACCGAAGCGTGGGCGCTC 6900
 1321 CTCAGCCAAATGATGCCCTGGCTTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATTG 1380
 6901 CTCAGCCAAATGATGCCCTGGCTTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATTG 6960
 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCTTGTAACTTTGTCTCTTCCAGAAAT 1440
 6961 CTACTCCTCTTTGGACCCCTGTATCTTTTACCTCTTGTAACTTTGTCTCTTCCAGAAAT 7020
 1441 GAAGCTGTAAAGCTACAGATGCTTCAAAATGGAACCCCA 1481
 7021 GAAGCTGTAAAGCTACAGATGCTTCAAAATGGAACCCCA 7061

RESULT 20

AAX25665

ID AAX25665 standard; cDNA to mRNA; 7582 BP.

XX AC AAX25665;

XX DT 21-MAY-1999 (first entry)

XX Complete human endogenous retrovirus W genome.

XX KW Clone; human endogenous retrovirus; genome; autoimmune disease;
 multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
 disseminated lupus erythematosus; pregnancy; chromosomal marker; 88.

XX OS Human endogenous retrovirus.

XX PN W09902696-AL.

XX PD 21-JAN-1999.

XX PF 06-JUL-1998; 98WO-PR001442.

XX PR 07-JUL-1997; 97FR-00008815.

XX PA (INMR) BIO MERIEUX.

XX PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;

XX PS WPI; 1999-120897/10.

XX PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
 exclusively in placenta and useful in diagnosis and therapy of autoimmune
 disease, and abnormal or failed pregnancy.

XX CC Claim 1; Page 71-74; 106pp; French.

XX CC This sequence represents the complete sequence of the human endogenous
 retrovirus (HERV) W genome. The nucleic acids, their fragments or
 peptides encoded by them are markers of autoimmune disease (e.g. multiple
 sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
 insulin-dependent diabetes and related pathologies) and of abnormal or
 unsuccessful pregnancy and can be used as chromosomal markers for
 susceptibility to these conditions, or proximity markers of genes
 associated with this susceptibility

XX SQ Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;

Query Match 89.3%; Score 1322.6; DB 2; Length 7582;

Best Local Similarity 92.1%; Pred. No. 0;

Matches 1364; Conservative 24; Mismatches 93; Indels 0; Gaps 0;

1 ATGCGCCCTCCCTATCATATCTTTCTTTACTGTTCTTCTTACCCCTTTTGGCTCTCACT 60

5581 ATGCGCCCTCCCTATCATATCTTTCTTCTKTASTGTTTYYACCTSTTTTCACTCTCACT 5640

QY 61 GCACCCCTTCATGCTGTGTATACCAACAGTAGCTCCCTTACCAAGAGTTTCTATGAGA 120
Db 5641 GCACCCCTTCATGCTGTGTATGACAGTAGCTCCCTYACCMAGAGTTTCTATGGAGA 5700
QY 121 ACGCGCTTCCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAAGGAAACTCC 180
Db 5701 ATGAGAGCTCCGGAAATATTGATGCCCATCGTATAGGAGTTTSTAAAGGAACCCCC 5760
QY 181 ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCGCACTCTTTGCAATG 240
Db 5761 ACCTTCACCTGCCACACCCATATGCCCGCAACTGCTATACTCTGCCACTCTTTGCAATG 5820
QY 241 CATGCAATACTCATTTATGGACAGGAAATGATTAATCCCTAGTTGCTCGAGGACTT 300
Db 5821 CATGCAATACTCATTTATGGACAGGAAATGATTAATCCCTAGTTGCTCGAGGACTT 5880
QY 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCATATCCAGTATGCTGTGATGGGGGTGGAATT 360
Db 5881 GGAGTCACTGTCTGTTGGACTTACTTCAACCAACTGGTATGCTGTGATGGGGGTGGAATT 5940
QY 361 CAAGGTCAGGCAAGAGAAAACAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
Db 5941 CAAGATCAGGCAAGAGAAAACAATGTAAAGGAAGTAAATCTCCCAACTCAACCSGGGTACAT 6000
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db 6001 GGCACCTCTAGCCCTACAAAGGACTAGTATCTCTCAAACTACATGAACCCCTCCGTACC 6060
QY 481 CATACTCGCTGTGAGCCCTATTAAATACCAACCTCACTCGGCTCCATGAGTCTCAGCC 540
Db 6061 CATACTCGCTGTGAGCCCTATTAAATACCAACCTCACTCGGCTCCATGAGTCTCAGCC 6120
QY 541 CAAAAACCTACTAATCTGTTGGATGTCCTCCCTCGCACTTCAGGCAATACATTTCAATC 600
Db 6121 CAAAAACCTACTAATCTGTTGGATGTCCTCCCTGAACTTCAGGCCATATGTTTCAATC 6180
QY 601 CCGTGTCTGAAACAATGGAACAATCTCAGCAGAGAAATAAACACCACTTCGTTTTAGTA 660
Db 6181 CCGTGTACTGAACAATGGAACAATCTCAGCAGAGAAATAAACACCACTTCGTTTTAGTA 6240
QY 661 GGACCTCTGTTTTCCAATCTGGAATAAACCCATCTCAAACTCACTCGTGTGTAATTT 720
Db 6241 GGACCTCTGTTTTCCAATSTGGAATAAACCCATCTCAAACTCACTCGTGTGTAATTT 6300
QY 721 AGCAATACTATAGACACACACGCTCCCAATGATCAGTGGGTGAACCTCCACACGA 780
Db 6301 AGCAATACTATAGACACACCAACTCCCAATGATCAGTGGGTGAACCTCCACACAA 6360
QY 781 ATAGTCTGCTACCTCAGGAATATTTTTGCTGTGTACCTCAGCCTATCATTTGTTG 840
Db 6361 ATAGTCTGCTACCTCAGGAATATTTTTGCTGTGTACCTCAGCCTATCATTTGTTG 6420
QY 841 AATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTTAGTGCCCTATGACATCTAC 900
Db 6421 AATGGCTCTTCAAGATCTATGTGCTTCTCTCATTTCTTAGTGCCCTATGACATCTAC 6480
QY 901 ACTGAACAAGATTTATACATCATGTGCTACCTTAAGCCCAACAACAAGAGTACCCATT 960
Db 6481 ACTGAACAAGATTTATACATCATGTGCTACCTTAAGCCCGCAACAACAAGAGTACCCATT 6540
QY 961 CTTCTTTTGTATCAGACGAGGAGTCTAGGACAGCTAGGTACTGCGATTGGCAGTATC 1020
Db 6541 CTTCTTTTGTATAGGACGAGGAGTCTAGGTGCTAGGTACTGCGATTGGCAGTATC 6600
QY 1021 ACAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGCTGACATGGACAG 1080
Db 6601 ACAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAATGCTGACATGGACAG 6660
QY 1081 GTCACTGACTCCCTGGTCACTTCCAGAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTT 1140
Db 6661 GTCCGCCACTCCCTGGTCACTTCCAGAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTT 6720
QY 1141 CAAAAATCGAAGAGCTTTTAGACTTGCTAAACCGCCAAAAGAGGGGAACTGTTTATTTT 1200

Db 6721 CRAAATCGAAGAGCTTTAGACTYGTAAACGCTGARAGAGGGGAACTGTTTATTTT 6780
QY 1201 GGAGAAAGACGCTGTTTATTTATGTTAAATCAATCCAGAAATGTCACGTAGAAAGTTAAAGAA 1260
Db 6781 GGGAAAGAAATGCTGTTTATTTATGTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 6840
QY 1261 ATTGAGATCGAATACAAATGTAGACGAGAGAGCTTCAAAACACCGAACGCTGGGCTC 1320
Db 6841 ATTGAGATCGAATACAAATGTAGACGAGAGAGCTTCAAAACACCTGACCCCTGGGCTC 6900
QY 1321 CTCAGCCAAATGATGCTGCTGCTGCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 6901 CTCAGCCRAATGATGCTGCTGCTGCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 6960
QY 1381 TTACTCTCTTTGACCCCTGCTATCTTTAACTCTGTTAACTGTTTAACTGTTTCTTCCAGAAAT 1440
Db 6961 CTACTCTCTTTGACCCCTGCTATCTTTTACCTCTGTTTAACTGTTTCTTCCAGAAATC 7020
QY 1441 GAAGCTGTAAAGCTACAGATGCTTACAAATGGAACCCCA 1481
Db 7021 GAAGCTGTAAAGCTACAAATGGAACCCCAAGATGCAATGCA 7061
RESULT 21
AAH20069
ID AAH20069 standard; DNA; 2782 BP.
XX AAH20069;
AC AAH20069;
XX
DT 08-AUG-2001 (first entry)
XX
DE HERV-W envelope protein G encoding nucleic acid.
KW Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW envelope protein; multiple sclerosis-related superantigen; vaccine;
KW surface antigen; transmembrane; multiple sclerosis; neuroprotective;
KW antisense-therapy; autoimmune disorder; ds.
XX Human endogenous retrovirus.
XX
FH Key Location/Qualifiers
5'UTR 1..762
FT /*tag= a
FT 763..2379
FT /*tag= b
FT /product= "HERV-W envelope protein G"
FT /transl_except= (pos:790..792,aa:Phe)
FT /transl_except= (pos:793..795,aa:Thr)
FT /transl_except= (pos:812..814,aa:Leu)
FT /transl_except= (pos:818..820,aa:Ser)
FT /transl_except= (pos:862..864,aa:Tyr)
FT /transl_except= (pos:865..867,aa:Gln)
FT /transl_except= (pos:1174..1176,aa:Arg)
FT /transl_except= (pos:1441..1443,aa:Leu)
FT /transl_except= (pos:1903..1905,aa:Gln)
FT /transl_except= (pos:2017..2019,aa:Lys)
FT /transl_except= (pos:2026..2028,aa:Arg)
FT /transl_except= (pos:2044..2046,aa:Arg)
FT /transl_except= (pos:2089..2091,aa:Gln)
FT /transl_except= (pos:2170..2172,aa:Asn)
FT 2380..2782
FT /*tag= c
XX
XX WO200131021-A1.
XX
XX 03-MAY-2001.
XX
XX 30-OCT-2000; 2000MO-BF010659.
XX
XX 28-OCT-1999; 99EP-00402690.
XX
XX (UYGE-) UNIV GENEVE.
PA

XX Conrad B, Mach B;
 XX WPI; 2001-316336/33.
 DR P-PSDB; AAB75138.
 XX New human retrovirus HERV-W ENV proteins/peptides having superantigen
 PT activity useful for diagnosing and treating multiple sclerosis.
 XX Claim 13; Fig 9; 94pp; English.
 XX On the basis of the PBS t-RNA motif used for the classification of human
 CC endogenous retrovirus (HERVs) the full length endogenous provirus which
 CC was located on the long arm of human chromosome 7 (7q21-22) has been
 CC designated HERV-W. The present invention describes proteins or peptides
 CC (I) having superantigen (SAG) activity comprising the ENV protein (ENV)
 CC of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
 CC have neuroprotective activity, and can be used in: vaccines; antisense-
 CC therapy; and HERV-W SAG active-inhibitors. (I) and encoding DNA/RNA are
 CC useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
 CC disorders. (I) are also useful for identifying substances (and optionally
 CC recovering) capable of binding to a retroviral superantigen associated
 CC with MS, substances capable of blocking SAG activity and substances
 CC capable of blocking transcription or translation of HERV-W retroviral
 CC superantigen. A protein or peptide derived from (I), modified to be
 CC devoid of SAG activity and being capable of generating an immune response
 CC against HERV-W retroviral SAG is useful in therapy. Nucleic acid
 CC molecules encoding (I) are useful as vaccines against MS. Substances
 CC capable of blocking SAG activity, capable of binding to a retroviral
 CC superantigen associated with MS, or capable of blocking transcription or
 CC translation of HERV-W retroviral superantigen for use in treating or
 CC preventing MS, obtained using (I) are useful for the treatment and
 CC prevention of MS. (I) and nucleic acids encoding them are useful for
 CC diagnosing autoimmune disease. The present sequence encodes the
 CC specifically claimed envelope protein of HERV-W designated G
 XX Sequence 2782 BP; 741 A; 768 C; 564 G; 709 T; 0 U; 0 Other;
 Query Match 88.3%; Score 1308.2; DB 5; Length 2782;
 Best Local Similarity 92.7%; Pred. No. 0;
 Matches 1373; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
 QY 1 ATGCGCCCTCCCTTATACATCTTTCTCTTACTGCTCTTACCCCTTTGCGCTCTCACT 60
 DB 763 ATGCGCCCTCCCTTATACATCTTTCTCTGAGTGTCTTTTACCCCTTTTCACTCTCACT 822
 QY 61 GCACCCCTCCATGCTGCTGATCAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
 DB 823 GCACCCCTCCATGCGCTGTATGACAGTAGCTCCCTCACCAGAGTTTCTATGGAGA 882
 QY 121 ACGCGGCTTCTGGAATATTGATGCGCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180
 DB 883 ATGACAGGCTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAACTCC 942
 QY 181 ACCTTCACTGCGCACCCATATGCGCGCACTGCTATTAATCTGTCAGTCTTTTGCATG 240
 DB 943 ACCTTCACTGCGCACCCATATGCGCGCACTGCTATTAATCTGTCAGTCTTTTGCATG 1002
 QY 241 CATGCAATATCTATTATTGACAGGGGAAATGATTAATCTAGTGTCTGAGGAGCTT 300
 DB 1003 CATGCAATATCTATTATTGACAGGGGAAATGATTAATCTAGTGTCTGAGGAGCTT 1062
 QY 301 GGAGCCATGCTGTTGACCTTACTTACCCCATACAGATATGCTGATGGGGTGGAAATT 360
 DB 1063 GGAGTCACTGCTGTTGACCTTACTTACCCCATACAGATATGCTGATGGGGTGGAGTT 1122
 QY 361 CAAGGTCAAGCAAGGAAACAGTAAAGCAAGCAATCTCCCACTGACCCCGGGGACAT 420
 DB 1123 CAAGGTCAAGCAAGGAAACAGTAAAGCAAGCAATCTCCCACTGACCCCGGGGACAT 1182
 QY 421 AGCACCCCTAGCCCTACAAAGGAGTACTTCTCTCAAACTACATGAACCCCTCCGTACC 480
 DB 1183 GGCACCTTAGCCCTACAAAGGAGTACTTCTCTCAAACTACATGAACCCCTCCGTACC 1242

QY 481 CATACTCGCTGCTGAGCCTATTTTATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC 540
 DB 1243 CATACTCGCTGCTGAGCCTATTTTATACACCCCTCACTCGGCTCCATGAGGTCTCGGCC 1302
 QY 541 CAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGCACTTCAGGGCATACATTTCAATC 600
 DB 1303 CAAACCCCTACTAACTGTTGGATGCTCCCTCCCTGAACTTCAGGGCATATGTTTCAATC 1362
 QY 601 CCGTGTCTGAAACAATGGAACAACCTTCAGACAGAAATAAACACCACTTCCTGTTTAGTA 660
 DB 1363 CCGTGTACCTGAAACAATGGAACAACCTTCAGACAGAAATAAACACCACTTCCTGTTTAGTA 1422
 QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCACTCACTCAAACTCACTGTGTAAATTTT 720
 DB 1423 GGACCTCTGTTTCCAACTCTGGAATAAACCACTCACTCAAACTCACTGTGTAAATTTT 1482
 QY 721 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
 DB 1483 AGCAATACTATAGACACACACAGCTCCCAATGATCAGGTGGGTAACTCTCTCCACACAA 1542
 QY 781 ATAGTCTGCTCCCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTTG 840
 DB 1543 ATAGTCTGCTCCCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTTG 1602
 QY 841 AATGGCTCTTCAAGAACTATGCTCTCTCTCAATTTTGTAGTGGCCCTATGACCATCTAC 900
 DB 1603 AATGGCTCTTCAAGAACTATGCTCTCTCTCAATTTTGTAGTGGCCCTATGACCATCTAC 1662
 QY 901 ACTGAACAAGATTTTATCAATCATGCTGTAACCTTAAGCCCAACAAACAAAGATACCCATT 960
 DB 1663 ACTGAACAAGATTTTATACAGTTATGTATCATATCTAAGCCCGCAACAAACAAAGATACCCATT 1722
 QY 961 CTTTCCTTTGTTATCAGACAGGAGTCTAGGAGAGTCTAGTCTGTCATTTGGCAGTATC 1020
 DB 1723 CTTTCCTTTGTTATAGAGCAGGAGTCTAGTCTGTCATTTGGCAGTATCCTGCGGTATC 1782
 QY 1021 ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGAACGAAACAG 1080
 DB 1783 ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGGACATGGAACGG 1842
 QY 1081 GTCACTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 DB 1843 GTCCGCGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1902
 QY 1141 CAAANTCGAAGCTTTAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 DB 1903 CGAAATCGAAGAGCTTTAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1962
 QY 1201 GGAGAAGAACGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1260
 DB 1963 GGGGAGNATGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2022
 QY 1261 ATTTCGAGATCGAATATCAATGTAGACAGAGAGCTTCAAAACACCGAAGCGTGGGGCTC 1320
 DB 2023 ATTTCGAGATCGAATATCAACGATATAGCAGAGAGCTTCAAAACACCGAAGCGTGGGGCTC 2082
 QY 1321 CTACGCCAATGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 DB 2083 CTCAGCCGATGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2142
 QY 1381 TTACTCTCTTTTGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 DB 2143 CTACTCTCTTTTGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2202
 QY 1441 GAAGCTGTAAAGCTTACAGATGGTCTTTTAAATGGAACCCCA 1481
 DB 2203 GAAGCTGTAAAGCTTACAGATGGTCTTTTAAATGGAACCCCA 2243

RESULT 22
 AAX25661
 ID AAX25661 standard; cDNA to mRNA; 2782 BP.

XX AC	AAX25661;	Db	1123	CAAGATCAGGCAAGAAACATGTAAGAAGTAATCTCCCACTCACCGGGTACAT	1182
XX DT	21-MAY-1999 (first entry)	Qy	421	AGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAAAAATAATAGAAACCTCGGTACC	480
XX DE	Human endogenous retrovirus W clone cl.PH74.	Db	1183	GGCACCTTAGCCCTACAAAGGACTAGTCTCTCAAAAATAATAGAAACCTCGGTACC	1242
XX KW	Clone; human endogenous retrovirus; genome; autoimmune disease; multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes; disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.	Qy	481	CATACCTGCCTGGTGGCTTATTTAATACCAACCTCCTCGGCTCCATAGAGTCTCAGCC	540
XX KW	Human endogenous retrovirus.	Db	1243	CATACCTGCCTGGTGGCTTATTTAATACCAACCTCCTCGGCTCCATAGAGTCTCAGCC	1302
XX OS	WO902696-A1.	Qy	541	CAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAGGCCCATACATTTCAATC	600
XX PN	21-JAN-1999.	Db	1303	CAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAGGCCCATATGTTTCAATC	1362
XX PD	06-JUL-1998; 98WO-FR001442.	Qy	601	CCTGTTCTGAAACAATGGAAACAACCTTCAGGACAGAAATAACACCACTTCGGTCTTAGTA	660
XX PF	07-JUL-1997; 97FR-00008815.	Db	1363	CCTGTACCTGAACAATGGAAACAACCTTCAGGACAGAAATAACACCACTTCGGTCTTAGTA	1422
XX PR	(INMR) BIO MERIEUX.	Qy	661	GGACCTCTGTTTCCAAATCTGAAATAACCCATACCTCAAAACCTCACCTGTGTAAATTT	720
XX PA	Besene F, Blond J, Bouton O, Mandrand B, Mallet F;	Db	1423	GGACCTCTGTTTCCAAATCTGAAATAACCCATACCTCAAAACCTCACCTGTGTAAATTT	1482
XX PI	WPI; 1999-120897/10.	Qy	721	AGCAATCTATAGACACCAACCACTCCCAATGCATCAGGTGGGTAAACCTCCCAACAGA	780
XX DR	New nucleic acid sequences from human endogenous retrovirus-W - expressed exclusively in placenta and useful in diagnosis and therapy of autoimmune disease, and abnormal or failed pregnancy.	Db	1483	AGCAATCTATAGACACCAACCACTCCCAATGCATCAGGTGGGTAAACCTCCCAACAGA	1542
XX PT	Claim 1; Page 60-63; 106pp; French.	Qy	781	ATAGTCTCCCTACCTCAGGAATATTTTGTCTGTGTGTAGTCTCAGCCTATCATTTGTTG	840
XX PP	This sequence represents clone cl.PH74 of the human endogenous retrovirus (HERV) W genome. The nucleic acids, their fragments or peptides encoded by them are markers of autoimmune disease (e.g. multiple sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-dependent diabetes and related pathologies) and of abnormal or unsuccessful pregnancy and can be used as chromosomal markers for susceptibility to these conditions, or proximity markers of genes associated with this susceptibility	Db	1543	ATAGTCTCCCTACCTCAGGAATATTTTGTCTGTGTGTAGTCTCAGCCTATCATTTGTTG	1602
XX PS	Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;	Qy	841	AATGGCTCTCAGGAATCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC	900
XX SQ	Query Match 88.2%; Score 1306.6; DB 2; Length 2782; Best Local Similarity 92.6%; Pred. No. 0; Matches 1372; Conservative 0; Mismatches 109; Indels 0; Gaps 0;	Db	1603	AATGGCTCTCAGGAATCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAC	1662
Qy	1 ATGGCCCTCCCTATCATACTTTTCTTTTACTGTTCTCTTACCCCTTTTCGCTCTCACT	Qy	901	ACTGAACCAAGATTTTATCAATCATGTCTGTAAGGCCCAACAAAGAGTACCCATT	960
Db	763 ATGGCCCTCCCTATCATATTTTCTCTGTAGTGTCTTTTACCCCTTTTCTCTCACT	Db	1663	ACTGAACCAAGATTTTATCAATCATGTCTGTAAGGCCCAACAAAGAGTACCCATT	1722
Qy	61 GCACCCCTCCCTGCTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA	Qy	961	CTTCTCTTTGTTATCAGAGCAGGAGTGTAGGACAGACTAGGTACTGGCATTTGGCAGTATC	1020
Db	823 GCACCCCTCCCTGCTGTATGACCAAGTAGCTCCCTCACCAGAGTTTCTATGGAGA	Db	1723	CTTCTCTTTGTTATGAGGACAGGAGTGTAGTGTGACACTAGGTACTGGCATTTGGCAGTATC	1782
Qy	121 ACGCGGCTTCTGGAATATTTATGATGCCCCATCATATAGGAGTTTATCTAAGGAACTCC	Qy	1021	ACAACTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG	1080
Db	883 ATGCAGCGTCCCGGAATATTTATGATGCCCCATCGTATAGGAGTCTTTCTAAGGAAACCC	Db	1783	ACAACTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAG	1842
Qy	181 ACCTTCACTGCCACACCATATGCCCCGCACTGTCTATAACTCTGCACTCTTTGCAATG	Qy	1081	GTCACTGACTCCCTGGTCACTTGTGCAAGATCAACTTAATCTCCTAGCAGCAGTACTCTTT	1140
Db	943 ACCTTCACTGCCACACCATATGCCCCGCACTGTCTATAACTCTGCACTCTTTGCAATG	Db	1843	GTCCGCGACTCCCTGGTCACTTGTGCAAGATCAACTTAATCTCCTAGCAGCAGTACTCTTT	1902
Qy	241 CATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAGTTGTCTCTGAGGACTT	Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGGAACCTGTTTATTTTATA	1200
Db	1003 CATGCAAAATCTCATTTATGGACAGGAAATGATTAATCCTAGTTGTCTCTGAGGACTT	Db	1903	CGAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGGAACCTGTTTATTTTATA	1962
Qy	301 GGAGCCACTGTCTGTGGACTTACTTACCCATACCAAGTATGCTGATGGGGTGAATT	Qy	1201	GGAGAGAAACGCTGTTTATTTATTTATTAATCAATCAGAAATGTTCACTGAGAAAGTTAAAGAA	1260
Db	1063 GGAGTCACTGTCTGTGGACTTACTTACCCAACTGGTATGTCTGATGGGGTGGAGTT	Db	1963	GGGGAAGAAATGCTGTTTATTTATTTATTAATCAATCAGAAATGTTCACTGAGAAAGTTAAAGAA	2022
Qy	361 CAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCACTGACCCCGGGGACAT	Qy	1261	ATTGAGATCGAATCAATGTTAGAGCAGGAGCTTCAAAACACGAAACGCTGGGGCCTC	1320
Db		Db	2023	ATTGAGATCGAATCAATGTTAGAGCAGGAGCTTCAAAACACGAAACGCTGGGGCCTC	2082
Qy		Qy	1321	CTCAGCAATGATGCTCTGGGTCTCTCCCTTTCTTAGGACCTCTAGCAGCTCTTAATATTG	1380
Db		Db	2083	CTCAGCAATGATGCTCTGGGTCTCTCCCTTTCTTAGGACCTCTAGCAGCTCTTAATATTG	2142
Qy		Qy	1381	TTACTCTCTTTGGACCTCTGATCTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGAAAT	1440
Db		Db	2143	CTACTCTCTTTGGACCTCTGATCTTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGAAAT	2202
Qy		Qy	1441	GAAGCTGTAAAGCTACAGATGCTTCAAAATGGAAACCCCA 1481	

Db	2203	GAAGCTGTGAACTCAAAATGGAGCCCAAGATGAGTCCAA	2243	Db	1003	CATGCAATACTACTATTATTGGACAGGAAAAATGATTAACTCCTAGTTGTCTCGGAGACTT	1062
RESULT 23				Qy	301	GGAGCCACTGTCTGTTGGACTTACTTACCCCAATACCAGTATGTCTGATGGGGTGGAAAT	360
AAA59211				Db	1063	GGAGTCACGTCTGTTGGACTTACTTACCCCAATAGTATGTCTGATGGGGTGGAGTT	1122
AC	AA59211			Qy	361	CAAGGTGAGCAAGAGAAAAACAAGTAAAGAGCAATCTCCCAATGACGCCGGGAGCAT	420
XX				Db	1123	CAAGATCAGCAAGAGAAAAACAATGTAAGAAAGTAACTCTCCCAATCTCACCGGGTACAT	1182
DT	07-NOV-2000	(first entry)		Qy	421	AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATCAATCAAACTCCGTPACC	480
XX				Db	1183	GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATCAAACTCCGTPACC	1242
KW	Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;			Qy	481	CATACCTCGCTGCTGAGCCCTATTAAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC	540
KW	gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.			Db	1243	CATACCTCGCTGCTGAGCCCTATTAAATACACCCCTCACTCGGCTCCATGAGGTCTCAGCC	1302
XX	Homo sapiens.			Qy	541	CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCGCATCTCAGGCCATACATTTCAATC	600
XX	WO200043521-A2.			Db	1303	CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCGCATCTCAGGCCATACATTTCAATC	1362
PD	27-JUL-2000.			Qy	601	CCTGTTCTGAAACAATGGAAACAATTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA	660
XX	21-JAN-2000; 2000WO-FR000144.			Db	1363	CCTGTTCTGAAACAATGGAAACAATTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA	1422
XX	21-JAN-1999; 99FR-00000888.			Qy	661	GGACCTCTGTTTCCCAATCTGGAATAAACCCATACCTCAAACTCCTGCTGTAAATTT	720
XX	(INMR) BIO MERIEUX.			Db	1423	GGACCTCTGTTTCCCAATCTGGAATAAACCCATACCTCAAACTCCTGCTGTAAATTT	1482
PA	Paranhos-Baccala G, Mallet F, Voisset C;			Qy	721	AGCAATACTATAGACAAACACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA	780
PI	WPI; 2000-499229/44.			Db	1483	AGCAATACTATAGACAAACACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA	1542
DR	New nucleic acid from human endogenous retrovirus, useful e.g. for			Qy	781	ATAGTCTGCTACCTACCTCAGGAATATTTTGTCTGTTGTTGCTCAGCTCATCTGTTTG	840
PT	diagnosis of autoimmune disease and complications of pregnancy, contains			Db	1543	ATAGTCTGCTACCTACCTCAGGAATATTTTGTCTGTTGTTGCTCAGCTCATCTGTTTG	1602
PT	at least part of the gag gene.			Qy	841	AATGGCTCTTTCAGAACTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	900
XX	Disclosure; Page 46-47; 53pp; French.			Db	1603	AATGGCTCTTTCAGAACTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1662
XX	The present sequence represents an endogenous retroviral nucleic acid			Qy	901	ACTGAACAAGATTTATACAACTATGTCGTACCTAAGCCCAACCAAAAGAGTACCCATT	960
CC	fragment, which is associated with an autoimmune disease, and is			Db	1663	ACTGAACAAGATTTATACAGTTATGTCATATCTAAGCCCAACCAAAAGAGTACCCATT	1722
CC	integrated into the human genome. The fragment is originally derived from			Qy	961	CTTCTCTTTGTTATCAGACGAGGTGCTAGGACAGACTAGGTACTGCGCATGGCAGTATC	1020
CC	a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W			Db	1723	CTTCTCTTTGTTATAGGACGAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGTGCTAGGTG	1782
CC	retrovirus is associated with autoimmune disease, failure of pregnancy or			Qy	1021	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATTAATGCTGACATGGAACAG	1080
CC	disorders of pregnancy. The nucleic acid fragment, or proteins derived			Db	1783	ACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATTAATGCTGGAACAGG	1842
CC	from it, are useful for diagnosis of autoimmune disease (specifically			Qy	1081	GTCACTGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGAGTACTCTT	1140
CC	multiple sclerosis) and for monitoring pregnancy. The nucleic acid			Db	1843	GTCCGCGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGAGTACTCTT	1902
CC	fragments may also be used for in situ labelling of isolated chromosomes,			Qy	1141	CAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTTTATTTTA	1200
CC	while the transcription product can be used to study or monitor T cell			Db	1903	CGAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAGAGAGGGGAACTGTTTATTTTA	1962
CC	proliferation in vitro			Qy	1201	GGAGAACACGCTGTTTATTTATGTTTAACTCAAGAAATGTCCTGAGAAAGTAAAGAA	1260
XX	Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;			Db	1963	GGGGAAGAAATGCTGTTTATTTATGTTTAACTCAAGAAATGTCCTGAGAAAGTAAAGAA	2022
Qy	Query Match 88.2%; Score 1306.6; DB 3; Length 2782;			Qy	1261	ATTCGAGATCGAATACAAATGTAGACGAGAGAGGTTCAAAACACCGAACCGTGGGGCTC	1320
Db	Best Local Similarity 92.6%; Pred. No. 0;			Db	2023	ATTCGAGATCGAATACAAATGTAGACGAGAGAGGTTCGAAACACTGGACCTGGGGCTC	2082
	Matches 1372; Conservative 0; Mismatches 109; Indels 0; Gaps 0;			Qy	1321	CTCAGCCAATGGATGCGCTGGGTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
1	ATGGCCCTCCCTTATCATCTTTCTTTCTTACTGTTCTTTACCCCTTTCGCTCTCACT	60					
763	ATGGGCTCCCTTATCATATTTTCTGTTAGTGTCTTTCCACCTGTTTCACTCTCACT	822					
61	GCACCCCTCCATGCTGTGACAAACAGTAGTCTCCCTTACCAAGAGTCTTCTATGAAGA	120					
823	GCACCCCTCCATGCGGTGTATGACAGTAGTCTCCCTACCCAGAGTCTTCTATGGAGA	882					
121	ACGGGGCTTCTCGAAATATTGATGCCCATCATATAGGAGTCTTCTAAGGGAAATCC	180					
883	ATGCAGCGTCCCGGAAATATTGATGCCCATCGTATAGGAGTCTTCTAAGGGAAATCC	942					
181	ACCTTCACTGCCACACCATATGCCCCGGAACCTGCTATACCTGCGCACTCTTTGCATG	240					
943	ACCTTCACTGCCACACCATATGCCCCGGAACCTGCTATACCTGCGCACTCTTTGCATG	1002					
241	CATGCAATACTCATTATTGGACAGGGAATGATTAACTCCTAGTTGTCTCGGAGGACTT	300					

Db 2083 CTGAGCGATGGATGCGCTGGATTCCTCCCTCTTAGGACCTCTAGCAGCTAATAATTG 2144

Qy 1381 TTACTCCTCTTTGGAGACCTGTATCTTTAACTCCTCTTGTTAAGTTTGTCTTCCAGAATT 1440

Db 2143 CTACTCCTCTTTGGAGACCTGTATCTTTGACCTCCTTGTAACTTGTCTCTTCCAGAATC 2202

Qy 1441 GAAGCTGTAAGCTACAGATGGTCTTACAATGGAAACCCCA 1481

Db 2203 GAAGCTGTGAAGACTACAAATGGAGCCCAAGATGCAATCCAA 2243

RESULT 24

ADE09587

ID ADE09587 standard; DNA; 3464 BP.

XX AC ADE09587;

XX 29-JAN-2004 (first entry)

DE Novel DNA-related contig nucleotide sequence #309.

XX novel gene; novel protein; tissue marker; molecular weight marker;

XX chromosome marker; genetic disorder; contig; ds.

KW Unidentified.

OS

XX WO2003054152-A2.

XX 03-JUL-2003.

XX 10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 14-MAR-2002; 2002US-0365384P.

PR 12-APR-2002; 2002US-0372381P.

PR 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.

XX

FA (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;

PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

XX WPI; 2003-569235/53.

DR

XX New polynucleotides, useful for expressing recombinant proteins for

PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.

XX

XX Disclosure; SEQ ID NO 2131; 1177pp; English.

XX

CC The invention comprises the amino acid and coding sequences of novel

CC proteins. The DNA and protein sequences of the invention are useful as:

CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or

CC tags; to identify chromosomes or to map related gene positions; and to

CC compare with endogenous DNA sequences in patients to identify potential

CC genetic disorders. The present DNA sequence was used in the

XX exemplification of the invention.

XX

XX Sequence 3464 BP; 1038 A; 913 C; 666 G; 845 T; 0 U; 2 Other;

Qy Query Match 87.7%; Score 1298.2; DB 10; Length 3464;

Best Local Similarity 93.0%; Pred. No. 0;

Matches 1383; Conservative 0; Mismatches 98; Indels 6; Gaps 2;

Qy 1 ATGGCCCTCCCTTATCATACTTTCTTTACTGTCTTACCCCTTCGCTCTCACT 60

Db 1430 ATGGCCCTCCCTTATCATATTCTTTCTTTACTGTCTTTACCTCTTCACTCTCACT 1489

1135 GTCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAGAGGGGAACCTGTGTTA 1194
 2570 GTCTTCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGGGGAACCTGTGTTA 2629
 1195 TTTTGGAGAGAGAGCTGTTATATGTTATGTTATCAATCAATCGAATTTGCTACTGAGAAGTT 1254
 2630 CTTTGGAGAGAGAGCTGTTATATGTTATGTTATCAATCAATCGAATTTGCTACTGAGAAGTT 2689
 1255 AAGAAATTCGAGATCGAATACAAATGAGAGAGAGGCTTCAAAACACCGGAACGCTGG 1314
 2690 AAGAAATTCGAGATCGAATACAAATGAGAGAGAGGCTTCAAAACACCGGAACGCTGG 2749
 1315 GGCTCTCTCAGCAATGGATGCGCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTA 1374
 2750 GGCTCTCTCAGCAATGGATGCGCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTATA 2809
 1375 ATATGTTACTCTCTTTGGACCTGTAATCTTTAACTCTCTGTTAAGTTTGTCTCTTC 1434
 2810 ATATGTTACTCTCTTTGGACCTGTAATCTTTAACTCTCTGTTAAGTTTGTCTCTTC 2869
 1435 AGAATTGAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 2870 AGAATTGAAGCTGTAAGCTACAAATGGAACCCCAAGATGAGTCCAA 2916

RESULT 25

ADFS9718 ID ADFS9718 standard; cDNA; 9502 BP.

AC ADF59718;

XX ADF59718;

XX 12-FEB-2004 (first entry)

XX Human contig polynucleotide sequence SEQ ID NO:2085.

XX biological activity; genetic engineering; hybridisation probe; oligomer;

XX primer; chromosome mapping; gene mapping; recombinant protein production;

XX human; gene; ss.

XX Homo sapiens.

XX WO2003080795-A2.

XX 02-OCT-2003.

XX 09-AUG-2002; 2002WO-US025485.

XX 09-AUG-2001; 2001US-0311261P.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX WPI; 2003-876918/81.

XX P-PSDB; ADF60170.

XX New polynucleotides, useful as hybridization probes, oligomers or

XX primers, for chromosome or gene mapping, for the recombinant production

XX of proteins, and for generating antisense DNA or RNA.

XX Example 2; SEQ ID NO 2085; 571pp; English.

XX The present invention describes isolated polynucleotide sequences (I),

XX which encode polypeptides (II) with biological activity. Also described:

XX (1) a vector comprising (I); (2) an expression vector comprising (I); (3)

XX a host cell genetically engineered to comprise (I) which is operatively

XX associated with a regulatory sequence that modulates expression of (I) in

CC (I). The polynucleotides (I) can be used as hybridisation probes,
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant
 CC production of proteins, and for generating antisense DNA or RNA. The
 CC present sequence represents a human contig polynucleotide sequence, which
 CC is used in an example from the present invention.

XX Sequence 9502 BP; 2813 A; 2433 C; 1992 G; 2263 T; 0 U; 1 Other;

SQ Query Match 87.7%; Score 1298.2; DB 10; Length 9502;
 Best Local Similarity 93.0%; Pred. No. 0;
 Matches 1383; Conservative 0; Mismatches 98; Indels 6; Gaps 2;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTTCTTCTTACCCCTTTGCTCTCACT 60
 DB 1430 ATGGCCCTCCCTTATCATATTTTCTTTTACTGTTCTTTTACCCCTTTTCACTCTCACT 1489
 QY 61 GCACCCCTCCCTGCTGTACCAACAGTAGCTCCCTTTACCAAGAGTTTCTTATGAAGA 120
 DB 1490 GCACCCCTCCCTGCTGTATGACAGTAGCTCCCTTTACCAAGAGTTTCTTATGAAGA 1549
 QY 121 ACAGCGCTTCTTGGAAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCTC 180
 DB 1550 ATGACGGTCCCGGAATATTTGATGCCCATCGTATAGGAGTCTTTCTAAGGGAATCTC 1609
 QY 181 ACCTTCACTGCCACACCCCATATGCGCCGCACTGCTATCACTCTGCGACTTTTGCATG 240
 DB 1610 ACCTTCACTGCCACACCCCATATGCGCCGCACTGCTATCACTCTGCGACTTTTGCATG 1669
 QY 241 CATGGAATACTCATTTATGACAGGGAATCATTAATCTTAGTTGCTCTGAGGAGACTT 300
 DB 1670 CATGGAATACTCATTTATGACAGGGAATCATTAATCTTAGTTGCTCTGAGGAGACTT 1729
 QY 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCATACCACTATGCTGATGGGGTGGAAAT 360
 DB 1730 GGAGTCACTGTCTGCGACTTACTTCAACCATACCACTATGCTGATGGGGTGGAGTT 1789
 QY 361 CAAGGTGAGCAAGAGAAAACAAAGTAAGGAAGCAATCTCCCACTGACCCGGGAGACAT 420
 DB 1790 CAAGATCAGCAAGAGAAAACATGTAAAAGAGCAATCTCCCACTGACCCGGGAGACAT 1849
 QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCCCTGACC 480
 DB 1850 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCCCTGACC 1909
 QY 481 CATACCTCGCTGCTGAGCCCTATTTAATACCAACCTCTCACTCGGCTCCAATGAGGTCTCAGCC 540
 DB 1910 CATACCTCGCTGCTGAGCCCTATTTAATACCAACCTCTCACTCGGCTCCAATGAGGTCTCAGCC 1969
 QY 541 CAAAACCTACTAACTGTTGGATGTCCTCCCTGCACTTTCAGGCCATATCAATTCATC 600
 DB 1970 CAAAACCTACTAACTGTTGGATGTCCTCCCTGCACTTTCAGGCCATATGTTTCAATC 2029
 QY 601 CTTGTTCTGAAATGGAACAACTTCAGCAGAGAAATAACACCACTTCGTTTATGTA 660
 DB 2030 CTTGTTCTGAAATGGAACAACTTCAGCAGAGAAATAACACCACTTCGTTTATGTA 2089
 QY 661 GGACCTCTGTTTCCAACTCTGGAATAAACCACTACCTCAAACTCTGTGTAATAATTT 720
 DB 2090 GGACCTCTGTTTCCAACTCTGGAATAAACCACTACCTCAAACTCTGTGTAATAATTT 2149
 QY 721 AGCAATACTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
 DB 2150 AGCAATACTATAGACACAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 2209
 QY 781 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTTGACCTCAGCCTATCATTTGTTG 840
 DB 2210 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTTGACCTCAGCCTATCATTTGTTG 2269
 QY 841 AATGGCTCTTTCAGAACTATGTTGCTTCTCTCAATTTTAGTGGCCCCCTATGACCATCTAC 900
 DB 2270 AATGGCTCTTTCAGAACTATGTTGCTTCTCTCAATTTTAGTGGCCCCCTATGACCATCTAC 2329
 QY 901 ACTGAAACAAGATTTTATCAATCATGTGTC---GTACCTAAGCCCCCAACAAAGAGTACC 956

Db 2330 ACTGAACAGATTATACAGTTATGTCAATATCCTTAAGCCCCCGACAAAGAGTACC 2389
 QY 957 CATTTCTTCTTTTGTATCAGACGAGGAGTGTAGCGAGCTAGTACTGGCAATTGGCAG 1016
 Db 2390 CATTTCTTCTTTTGTATTTGGAGCAGGAGTACTAGGCGGACTAGTACTGGCAATTGGCGG 2449
 QY 1017 TATCACAACCTTACTCAGTTCTTACTACAACTATCTCAGAAATTAATGGTGACATGGA 1076
 Db 2450 TATTACAACCTTACTCAGTTCTTACTACCAAACTATCTCAAGAACTAAATGGTGACATGGA 2509
 QY 1077 ACAGGT--CACTGACTCCCTGGTCACTTGGCAAGTCACTTAACTCCCTAGCAGAGTA 1134
 Db 2510 ACAGGTGCGCCGACTCCCTGGTCACTTGGCAAGTCACTTAACTCCCTAGCAGAGTA 2569
 QY 1135 GTCTTTCAAAATCGAAGAGCTTTAGACTTTGCTTAAACCGCCAAAGAGGGGGAACCTGTTTA 1194
 Db 2570 GTCTTTCAAAATCGAAGAGCTTTAGACTTTGCTTAAACCGCTGAAAGAGGGGGAACCTGTTTA 2629
 QY 1195 TTTTATAGGAGAAGACGCTTTATATGTATTAATCAATCCAGAAATGTCACACTGAGAAAGTT 1254
 Db 2630 CTTTATAGGGAAGAATGCTTGTACTATGTATTAATCAATCCGGAATCGTCACTGAGAAAGTT 2689
 QY 1255 AAGAAATTCGAGATCGAATCAATGTAGACGAGGAGCTTCAAAACACCGAAGCGCTGG 1314
 Db 2690 AAGAAATTCGAGATCGAATCAACGTAGACGAGGAGCTTCGAAACACTGGACCTGG 2749
 QY 1315 GGCTCTCCTCAGCAATGGAATGACCTGGGTTCTCCCTTCTTAGGAGCTCTAGCAGCTCTA 1374
 Db 2750 GGCTCTCCTCAGCAATGGAATGACCTGGGTTCTCCCTTCTTAGGAGCTCTAGCAGCTATA 2809
 QY 1375 ATATTGTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCC 1434
 Db 2810 ATATTGTACTCTCTTTGGACCTGTATCTTTAACTCTCTGTTAAAGTTTGTCTCTCC 2869
 QY 1435 AGAATTCGAAGCTGTAAGAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 Db 2870 AGAATTCGAAGCTGTAAGAGCTACAAATGGAGCCCAAGATGCAGTCCAA 2916

RESULT 26

ID ABN97978 standard; DNA; 46340 BP.
 XX AC ABN97978;
 XX 01-AUG-2002 (first entry)
 XX Human retroviral sequence H13.
 XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
 XX multiple sclerosis; ds.
 XX Human endogenous retrovirus.
 XX WO9667395-A1.
 XX 29-DEC-1999.
 XX 23-JUN-1999; 99WO-FR001513.
 XX 23-JUN-1998; 98FR-00007920.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX Alliel PM, Perin J, Rieger F;
 XX WPI; 2000-160587/14.
 XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
 XX for diagnosis, treatment and prevention of autoimmune and neurological
 XX diseases.

PS Claim 15; Page 186-199; 225pp; French.
 XX The present invention relates to new nucleic acid sequences of human
 CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
 CC Regulatory elements associated with HERV-7q may alter expression of other
 CC genes (even remote genes) on the same chromosome, inducing immunological
 CC and/or neurological changes (which may be pathological or protective/
 CC curative). HERV-7q peptides can be used to improve efficiency of the
 CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
 CC sequences can be used in immunogenic or vaccinating compositions, for
 CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention
 XX Sequence 46340 BP; 16104 A; 8738 C; 8434 G; 13064 T; 0 U; 0 Other;
 SQ
 Query Match 86.6%; Score 1282; DB 3; Length 46340;
 Best Local Similarity 92.3%; Pred. No. 0;
 Matches 1395; Conservative 0; Mismatches 85; Indels 32; Gaps 3;
 QY 1 ATGGCCCTCCCTTATCATACTTTCTCTTACTGTCTTACTCCCTTTTGGCTCTCACT 60
 Db 31855 ATGGCCCTCCCTTATCATATTTTCTCTTACTGTCTTACTCCCTTTTCACTCTCACT 31914
 QY 61 GCACCCCTCCATGCTGTCAACACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
 Db 31915 GCACCCCTCCATGCTGTCAACACCAAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 31974
 QY 121 ACGCGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
 Db 31975 ATGGCGCTTCCAGAAAATATTGATGCCCATCAATAGGAGTTTACCTTAAAGGAAACTCC 32034
 QY 181 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATACTGCCACTCTTTTGCATG 240
 Db 32035 ACCTTCACTGCCACACCCATATGCCCGCACTGCTATACTGCCACTCTTTTGCATG 32094
 QY 241 CATGCAAAATCTCATTATTGGACAGGAAATGATTAACTCTAGTGTCTCGAGAGACTT 300
 Db 32095 CATGCAAAATCTCATTATTGGACAGGAAATGATTAACTCTAGTGTCTCGAGAGACTT 32154
 QY 301 GGAGCCACTGTCTGTGGACTTACTTACCCATACAGTATGTCTGTATGGGGGTGGAATT 360
 Db 32155 GGAGCCACTGTCTGTGGACTTACTTACCCATACAGTATGTCTGTATGGGGGTGGAATT 32214
 QY 361 CAAAGTCAAGGAGAGAAACAACTGAAAGAGCAATCTCCCACTGACCCGGGAGACAT 420
 Db 32215 CAAAGTCAAGGAGAGAGAAACAACTGAAAGAGCAATCTCCCACTGACCCGGGAGACAT 32274
 QY 421 AGCACCCTTAGCCCTTACAAAGGAGTACTGTCTCTCAAACTACATGAAACCTTCGTACC 480
 Db 32275 AGCACCCTTAGCCCTTACAAAGGAGTACTGTCTCTCAAACTACATGAAACCTTCGTACC 32334
 QY 481 CATACTGCGCTGTGAGCCTATTATTAATACCACTTCACTCGGCTCCATGAGGTCTCAGCC 540
 Db 32335 CATACTGCGCTGTGAGCCTATTATTAATACCACTTCACTCGGCTCCATGAGGTCTCAGCC 32394
 QY 541 CAAAACCTTACTAACTGTGTGGATGTCCTCCCTGCACTTCAGGCGCATACATTTCAATC 600
 Db 32395 CAAAACCTTACTAACTGTGTGGATGTCCTCCCTGCACTTCAGGCGCATACATTTCAATC 32454
 QY 601 CCTGTTTCTGAAACAATGGAACTTACAGCAGAGAAATAAACAACCTTCGTTTTAGTA 660
 Db 32455 CCTGTTTCTGAAACAATGGAACTTACAGCAGAGAAATAAACAACCTTCGTTTTAGTA 32514
 QY 661 GGACCTCTTGTTCNAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 720
 Db 32515 GGACCTCTTGTTCNAATCTGGAATAACCCATACCTCAAACTCACTGTGTAAATTT 32574
 QY 721 AGCAATACTATAGACAACACAGCTCCCAATGATGATCAGGTGGGTAAACCTCCACACGA 780
 Db 32575 AGCAATACTATAGACAACACAGCTCCCAATGATGATCAGGTGGGTAACTCTCTCCACACGA 32634

Db 1222 CAAACCCCTACTAATGTTGGTTGTGCTCCCTGTATTTTCAGGCCATGCAATTCATC 1163
Qy 601 CTGTTCTCTGAACAATGGAACAACCTTCAGCAAGAAATAAACAACATTCCTGTTTAGTA 660
Db 1162 CCGTACTCTGAACAATGGAACAACCTTCAGCAAGAAATAAACAACATTCCTGTTTAGTA 1103
Qy 661 GACCTCTGTTTCCAACTCTGAACAATGGAACAACCTTCAGCAAGAAATAAACAACATTC 720
Db 1102 GACCTCTGTTTCCAACTCTGAACAATGGAACAACCTTCAGCAAGAAATAAACAACATTC 1043
Qy 721 AGCAATCTATATAGACACAAACAGCTCCCAATGATCAGGTGGGTAAACCTCCCAACGA 780
Db 1042 AGCAATCTATATAGACACAAACAGCTCCCAATGATCAGGTGGGTAAACCTCCCAACGA 983
Qy 781 ATAGTCTGCTACCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 840
Db 982 ATAGTCTGCTACCTACCTCAGGAATATTTTTTGTCTGTGGTACCTCAGCCTATCATTTGTTG 923
Qy 841 AATGGCTCTCAGAACTATGCTTCCCTCTCATTTAGTCCCTATGACCTATGACCATCTAC 900
Db 922 AATGGCTCTCAGAACTATGCTTCCCTCTCATTTAGTCCCTCTCATTTAGTCCCTCTCATTTAG 864
Qy 901 ACTGAACAAGATTTATACAACTCATGCTGTAACCTTAAGCCCCCAACAACAAAGAGTACCCATT 960
Db 863 ACTGAACAAGATTTATACAACTCATGCTGTAACCTTAAGCCCCCAACAACAAAGAGTACCTATT 804
Qy 961 CTTCTCTTTGTTATCAGACAGAGTGTAGGAGACTAGGTACTGGGATTTGGGATGATC 1020
Db 803 CTTCTCTTTGTTATCGAGCAGAGTGTAGGAGACTAGGTACTGGGATTTGGGATGATC 744
Qy 1021 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAAC 1080
Db 743 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAATGG 684
Qy 1081 GTCACTGACTCCCTGCTGCTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
Db 683 GTTCCGACTCCCTGCTGCTTGAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 624
Qy 1141 CAAATCGAAGAGCTTTAGACTGCTAAACGCGCAAGAGGGGGAACCTGTTATTTTA 1200
Db 623 CAAATCGAAGAGCTTTAGACTGCTAAACGCGCAAGAGGGGGAACCTGTTATTTTA 564
Qy 1201 GGAGAGAGAGCTG-----TTATTATGTTAATCAA 1230
Db 563 GGGAAGAATGTTGTTATGTTATTTAGGGAAGAATGTTGTTATGTTAATCAA 504
Qy 1231 TCAGAAATGTCACCTGAGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAG 1290
Db 503 TCCTGAATGTCACAGAGAAAGTTGAAGAAATTCGAGATGGAATACAACGTAGAACAGAG 444
Qy 1291 GAGCTTC-AAAACACGAAGCTGGGGCTCCTCAGCAGATGATGATGCGCTGGTCTCCC 1349
Db 443 GAGCTTC-AAAACACGAAGCTGGGGCTCCTCAGCAGATGATGATGCGCTGGTCTCCC 384
Qy 1350 CTTCTTAGACCTCTAGAGCTCTAATATTGTTTACTCTCTTTGGACCTGTATCTTTAA 1409
Db 383 CTTCTTAGAGATCTAGAGCTCTAATATTGTTTACTCTCTTTGGACCTGTATCTTTAA 324
Qy 1410 CTTCTTGTGTTAGTTGTTCTCTTCCAGAAATGGAAGCTGTAAAGCTACAGATGCTTTACA 1469
Db 323 CTTCTTGTGTTAGTTGTTCTCTTCCAGAAATGGAAGCTGTAAAGCTACAGATGCTTTACA 264
Qy 1470 AATGGAACCCCA 1481
Db 263 AATGGAACCCCA 252

RESULT 28
ACN44334/c standard; DNA; 161334 BP.
XX ID ACN44334
AC ACN44334;

XX 18-NOV-2004 (first entry)
DT Human genomic sequence HCG32959.
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
KW Homo sapiens.
OS WO2003073826-A2.
PN 12-SEP-2003.
PD 28-FEB-2003; 2003WO-US006235.
PF 01-MAR-2002; 2002US-00087192.
PR (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 730; Opp; English.
XX The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: this patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published
XX SQ Sequence 161334 BP; 46474 A; 28689 C; 31549 G; 54602 T; 0 U; 20 Other;
Query Match 86.3%; Score 1277.4; DB 11; Length 161334;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 81; Indels 33; Gaps 4;
Qy 1 ATGGCCCTCCCTTATCATATCTTTCTTTTACTGTTCTTACCCCTTTCTCACT 60
Db 102037 ATGGCCCTCCCTTATCATATCTTTCTTTTACTGTTCTTACCCCTTTCTCACT 101978
Qy 61 GCACCCCTCCATGCTGTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 101977 GCACCCCTCCATGCTGTGTACACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 101918
Qy 121 ACGCGGCTTCTGGAATATTGATGCCCATCATATAGAGTTTATCTAAGGAAACTCC 180
Db 101917 ACGCGGCTTCTGGAATATTGATGCCCATCATATAGAGTTTATCTAAGGAAACTCC 101858
Qy 181 ACCTTCACTGCCACACACCATATGCGCCCAACTGCTATACTGTGCCACTCTTTGCATG 240
Db 101857 ACCTTCACTGCCACACACCATATGCGCCCAACTGCTATACTGTGCCACTCTTTGCATG 101798
Qy 241 CATGCAAACTACTATTATGGACAGGAAATGATTAATCTTAGTTGCTCTGAGGACTT 300
Db 101797 CATGCAAACTACTATTATGGACAGGAAATGATTAATCTTAGTTGCTCTGAGGACTT 101738
Qy 301 GGAGCCACTGCTGTTGGACTTTACTTACCCCATACAGTATGCTGATGGGGTGGAAAT 360
Db 101737 GGAGCCACTGCTGTTGGACTTTACTTACCCCATACAGTATGCTGATGGGGTGGAGTT 101678

QY 361 CAAGGTCAGGCAAGAGAAAAACAAGTAAAGAAAGCAATCTCCAACTGACCCCGGGACAT 420
 Db 101677 CAGATCAGGCAAGAGAAAAACAATGTAAGAGAGTAACTCTCCAACTGACCCCGGGTACAT 101618
 QY 421 AGCACCCCTAGCCCTACAAAGGAGTAGTTCTCTC-AAAACATACATGAACCCCTCCGTAC 479
 Db 101617 AGCACCCCTAGCCCTACAAAGGAGTAGATCTCTCTAAACATACATGAACCCCTCCATAC 101558
 QY 480 CCATACCTCGCTCGTGGAGCCCTATTTAATACCCCTCAGCTCAGCTCGGCTCCATGAGTCTCAGC 539
 Db 101557 CCATACCTCGCTCGTGGAGCCCTATTTAATACCCCTCAGCTCAGCTCGGCTCCATGAGTCTCGGC 101438
 QY 540 CCAAAACCCCTACTAACTGTTGGATGTGCTCTCCCTCCCTGCACTTCAGGCCATACATTTCAAT 599
 Db 101497 CCAAAACCCCTACTAACTGTTGGATGTGCTCTCCCTCCCTGCACTTCAGGCCATGCAATTTCAAT 101438
 QY 600 CCCTGTTCTGGAACAATGGAACAACCTTCAGCAGAGAAATAAACAACACCTTCCTGTTTGTAGT 659
 Db 101437 CCCTGTTCTGGAACAATGGAACAACCTTCAGCAGAGAAATAAACAACACCTTCCTGTTTGTAGT 101378
 QY 660 AGGACCTCTGTTTCCATCTGGAATAAACCCTACCTCCTAAACCTCAGCTGCTGTAATAAT 719
 Db 101377 AGGACCTCTGTTTCCATCTGGAATAAACCCTACCTCCTAAACCTCAGCTGCTGTAATAAT 101318
 QY 720 TAGCAATACATATAGACAAACCCAGCTCCCAATGATCAGCTGGGTAAACACCTCCCAACAG 779
 Db 101317 TAGCAATACATATAGACAAACCCAGCTCCCAATGATCAGCTGGGTAACTCTCCCAACAG 101258
 QY 780 AATAGTCTGCTACCTCAGCAATATTTTGTGTGTGATGCTGACCTCAGCTTATCATGTTT 839
 Db 101257 AATAGTCTGCTACCTCAGCAATATTTTGTGTGTGATGCTGACCTTATCATGTTT 101198
 QY 840 GAATGGCTCTCAGAACTATGCT 899
 Db 101197 GAATGGCTCTCAGAACTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 101139
 QY 900 CACTGAACAGATTTATACAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 959
 Db 101138 CACTGAACAGATTTATACAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 101079
 QY 960 TCTTCTCTTTTGTATACAGCAGAGTGTAGGAGATAGGTACTGGCAATGGCAGTAT 1019
 Db 101078 TCTTCTCTTTTGTATACAGCAGAGTGTAGGAGATAGGTACTGGCAATGGCAGTAT 101019
 QY 1020 CAAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACA 1079
 Db 101018 CAAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAATG 100959
 QY 1080 GGTCACTGACTCCCTGCTCAGCTGCAAGATCACTTAACTCCCTAGCAGCAGTAGTCTCT 1139
 Db 100958 GGTGCGGACTCCCTGCTCAGCTGCAAGATCACTTAACTCCCTAGCAGCAGTAGTCTCT 100899
 QY 1140 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGGAACCTGTTTATTTT 1199
 Db 100898 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGGAACCTGTTTATTTT 100839
 QY 1200 AGGAGAGAGAGCGTG-----TTATATGTTATATCA 1229
 Db 100838 AGGAGAGAGAGTGTGTTATATGTTATTTATGCGGAAGATGTTGTTATATGTTATATCA 100779
 QY 1230 ATCCAGAAATGTCTACTGAGAAGTTAAGAAATTCGAGATCGAATCAATCTAGAGCAGA 1289
 Db 100778 ATCCAGAAATGTCTACTGAGAAGTTAAGAAATTCGAGATCGAATCAATCTAGAGCAGA 100719
 QY 1290 GGAGCTTC-AAAACACCAAGCGCTGGGCGCTCTCAGCCAAATGATGCTCCCTGGGTTCTCC 1348
 Db 100718 GGAGCTTC-AAAACACCAAGCGCTGGGCGCTCTCAGCCAAATGATGCTCCCTGGGTTCTCC 100659
 QY 1349 CCTTCTTAGGACCTCTAGCAGCTCTAATATGTTTACTCTCTCTTTGGACCCCTGATCTTTA 1408
 Db 100658 CCTTCTTAGGATCTCTAGCAGCTCTAATATGTTTACTCTCTCTTTGGACCCCTGATCTTTA 100599
 QY 1409 ACCTCCTTGTAGTGTCTCTCTCCAGAAATGGAAGCTGTAAAGCTACAGATGCTCTTAC 1468

Db 100598 ACCTCTCTTGAAGTTGTCTCTTCCAGAAATCAAGTTGTAAAGCTACAAATCGTTCTTC 100539
 QY 1469 AAATGGAACCCCA 1481
 Db 100538 AAATGGAACCCCA 100526
 RESULT 29
 ABN97930
 ID ABN97930 standard; DNA; 2784 BP.
 XX AC ABN97930;
 XX XX
 DT 01-AUG-2002 (first entry)
 XX Human retroviral sequence HE2.
 XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
 KW multiple sclerosis; ds.
 XX Human endogenous retrovirus.
 XX OS
 XX PN WO967395-A1.
 XX PD 29-DEC-1999.
 XX PF 23-JUN-1999; 99WO-FR001513.
 XX PR 23-JUN-1998; 98FR-00007920.
 XX PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX PI Alliel PM, Perin J, Rieger F;
 XX WPI; 2000-160587/14.
 XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used
 PT for diagnosis, treatment and prevention of autoimmune and neurological
 PT diseases.
 XX Claim 3; Page 133; 225pp; French.
 XX The present invention relates to new nucleic acid sequences of human
 CC endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
 CC Regulatory elements associated with HERV-7q may alter expression of other
 CC genes (even remote genes) on the same chromosome, inducing immunological
 CC and/or neurological changes (which may be pathological or protective/
 CC curative). HERV-7q peptides can be used to improve efficiency of the
 CC immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
 CC sequences can be used in immunogenic or vaccinating compositions, for
 CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention
 XX SQ Sequence 2784 BP; 816 A; 755 C; 535 G; 678 T; 0 U; 0 Other;
 Query Match 83.7%; Score 1239; DB 3; Length 2784;
 Best Local Similarity 91.0%; Pred. No. 0;
 Matches 1364; Conservative 0; Mismatches 100; Indels 35; Gaps 3;
 QY 2 TGCGCCCTCCCTTATCATACTTTTCTCTTACTGTTCTTTACCCCTTTTCGCTCTCACTG 61
 Db 903 TGCGCCCTCCCTTATCATATAATTTTCTCTTACTGTTCTTTACCCCTTTTCACTCTCACTG 962
 QY 62 CACCCCTCCATCGCTGTGTAC-----AACCACTAGCTCCCTTAC 102
 Db 963 CACCCCTCCATCGCTTGCCTGACCCCGCTCCATGCCCGCTCCATGCTAGCTCCCTTAG 1022
 QY 103 CAAGAGTTTCTATGAAGAACCGGCTTCTCTGGAATAATTGATGCCCATCATATAGAGT 162

Db 1023 CAAGAGTTTCTATGAGAAATGAGCGTCCCGAAATATTTGATGCCCCATTGTTATAGGAGT 1082
Qy 163 TTATCTAAGGGAATCTCCACCTTCACTGCCCCACACCCATATATCCCGCGCAACTGCTATTAAC 222
Db 1083 TTATCTAAGGGAACCCCACTTCACTGCCCCACACCCATATATCCCGCGCAACTGCTATTAAC 1142
Qy 223 TCTGCCACTTTTGCATGATCAATATCTGATTTGACAGGGAATGATTAATCTCT 282
Db 1143 TCTGCCACTTTTGCATGATCAATATCTGATTTGACAGGGAATGATTAATCTCT 1202
Qy 283 AGTTGTCTGGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCCATACAGTATG 342
Db 1203 AGTTGTCTGGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCCATACAGTATG 1249
Qy 343 TCTGATGGGGTGGAAATCAAGGTGAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCC 402
Db 1250 TCTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACATATAAGGAAGTAAATCTCC 1309
Qy 403 CAACTGACCCGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAAAACTA 462
Db 1310 CAACTGACCTGGGTACATAGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAAAACTA 1369
Qy 463 CATGAACCCCTCGTACCCATATCTGCTGTGGCTATTTAAATACCAACCCCTCACTGG 522
Db 1370 CATGAACCCCTCGATACCCATATCTGCTGTGGCTATTTAAATACCAACCCCTGACTGG 1429
Qy 523 CTCCTAGAGGTCTGAGCCCAAAACCCCTACTAATCTGTTGATGTGCTCCCTCGACTTC 582
Db 1430 CTCCTAGAGGTCTGAGCCCAAAACCCCTACTAATCTGTTGATGTGCTCCCTCGACTTC 1489
Qy 583 AGGCCATACATTTCAATCCTGTCTGTAACATGGAACAACCTTCAGCACAGAAATAAAC 642
Db 1490 AGGCCATACATTTCAATCCTGTATACCTGAACAATGGAAACAATTCAGCACAGAAATAAAC 1549
Qy 643 ACCACTTCGGTTTGTAGTAGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAAAC 702
Db 1550 ACCACTTCGTTTGTAGTAGCTCTCTTCCAAATCTGGAATAACCCATACCTCAAAAC 1606
Qy 703 CTCACCTGTGTAAATTTAGCAATATATAGACACAACACAGCTCCCAATGCATCAGGTGG 762
Db 1607 CTCACCTGTGTAAATTTAGCAATATATAGACACAACACAGCTCCCAATGCATCAGGTGG 1666
Qy 763 GTAAACCTCCACACAGATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTACC 822
Db 1667 GTAACTTCCTCCACACAGATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGCTACC 1726
Qy 823 TCAGCCTATCATTTGTTGATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTG 882
Db 1727 TCAGCCTATCATTTGTTGATGGCTCTTCAGAACTATGTGCTTCTCTCATTTCTTAGTG 1786
Qy 883 CCCCCTATGACCATCTACACTGAACAAGATTTATACAATCATGTCTGCTACCTAAGCCCCAC 942
Db 1787 GCCCCTATGCCATCTACACTGAACAAGATTTATACAATCATGTCTAAGCCCCCG 1846
Qy 943 AACAAAGAGTACCCATCTCTCTTTGTTATCAGAGCAGGAGTGTAGCGAGTAGCT 1002
Db 1847 AACAAAGAGTACCCATCTCTCTTTGTTATTTGGAGCAGGAGTGTAGCGGAGTAGCT 1906
Qy 1003 ACTGGCATTTGGCAGTATCAAACTCTACTCAGTTCTTACTACAAATCTCTCAAGAAATA 1062
Db 1907 ACTGGCATTTGGCAGTATCAAACTCTACTCAGTTCTTACTACAAATCTCTCAAGAAATA 1966
Qy 1063 AATGGTGACATGAAACAGGTCTACTGCTCCCTGGTCACTTGAAGATCAATCACTACTCC 1122
Db 1967 AATGGTGACATGAAATGGGTGGTGTATACCTGGTCACTTGAAGATCAATCACTACTCC 2026
Qy 1123 CTAGCAGCAGTACTCTTCAAAATCGAAGCTTTAGACTTGTAAACCGCCAAAGAGGG 1182
Db 2027 CTAGCAGCAGTACTCTTCAAAATCGAAGCTTTAGACTTGTAAACCGCGGAAGCGGG 2086
Qy 1183 GGAACTGTTTATTTTATAGAGAGAAACGCTGTTTATTTATGTTAATCAATCCAGATTGTC 1242
Db 2087 GGAACTTTTTATTTTATAGAGAAATGCTGTTTATTTATGTTAATCAATCCGGAATCATC 2146

Qy 1243 ACTGAGAAAGTTAAAGAAATTCGAGATCGAATCAATGTAGAGCAGAGGAGCTTCAAAAC 1302
Db 2147 ACCGAGAAAGTTAAAGAAATTCAGGTGCAATATACGTAGAGCAAGAGAGCTGCAAAAC 2206
Qy 1303 ACCGAAACCTCGGGGCTCTCCTCAGCAATGGATGCCCTGGGTTCTCCCTTTTAGGACT 1362
Db 2207 ACTGGACCTCGGGGCTCTCCTCAGCAATGGATGCCCTGGATTCTCCCTTTTAGGACT 2266
Qy 1363 CTAGCAGCTTAATATTTGTTACTCTCTTTGGACCTGTATCTTTAACTCTCTCTTAAG 1422
Db 2267 CTAGCAGCTTAATATTTGTTACTCTCTTTGGACCTGTATCTTTAACTCTCTCTCTTAAG 2326
Qy 1423 TTTGTCTCTCCAGAAATTTGAAGCTGTAAGCTTACAGATGGTCTTACAAATGGAAACCCCA 1481
Db 2327 TTTGTCTCTTCCAGAAATTTGAAGCTGTAAGCTTACAGATGGTCTTCAATGGAGCCCCA 2385

RESULT 30
AA114608
ID AA114608 standard; DNA; 1894 BP.
XX
AC AA114608;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #4541 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.
XX
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX
PI WPI; 2001-488901/53.
XX
DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 4541; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 76.4%; Score 1131.6; DB 4; Length 1894;
Best Local Similarity 92.4%; Pred. No. 0;

Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;	
QY	164 TATCTAAGGAAACTCCACCTTCCAGTCCGACACCATATGCCCCGCAACTGTCTATAACT 223
Db	
QY	1 TATCTAAGGAAACCCCACTTCCAGTCCGACACCATATGCCCCGCAACTGTCTATAACT 60
Db	
QY	224 CTGCCACTCTTTTCATGTCATGCAAACTACTCATTTATTGGACAGGAAATGATTAATCCTTA 283
Db	
QY	61 CTGCCACTCTTTTCATGTCATGCAAACTACTCATTTATTGGACAGGAAATGATTAATCCTTA 120
Db	
QY	284 GTTGTCTGAGGACTTTGAGGACCACTGTCTGTTGGACTTACTTTCACCATACCACTATGT 343
Db	
QY	121 GTTGTCTGAGGACTTTGGA-----GGACTCACTTCACTCATACCACTATGT 167
Db	
QY	344 CTGATGGGGTGGAAATTCAGGTCAGGCAAGAGAAACAAAGTAAAGAGAGCAATCTCCC 403
Db	
QY	168 CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAACACATPAAGGAAGTAAATCTCCC 227
Db	
QY	404 AACTGACCCGGGACATAGCACCCCTAGCCCTTACAAAGGACTAGTTTCTCTCAAAACTAC 463
Db	
QY	228 AACTGACCTGGGTACATAGCACCCCTGGCCCTTACAAAGGACTAGATCTCTCAAAACTAC 287
Db	
QY	464 ATGAAACCCCTCCGTACCCATCTAGCTGCGCTGTGAGCCTTATTATACACCCCTCACTGGC 523
Db	
QY	288 ATGAAACCCCTCCATACCCATCTAGCTGCGCTGTGAGCCTTATTATACACCCCTCACTGGC 347
Db	
QY	524 TCCATGAGGTCTGAGCCCAACCCCTACTAATCTGTTGGATGCTCCCTCCCTGCACCTCA 583
Db	
QY	348 TCCATGAGGTCTGAGCCCAACCCCTACTAATCTGTTGGATGCTCCCTCCCTGCACCTTCA 407
Db	
QY	584 GGCATACATTTCAATCCCTGTTCTCGAACAAATGGAACAACTTTCAGCACAGAAATAAACA 643
Db	
QY	408 GGCATACATTTCAATCCCTTATCTGAAACAAATGGAACAACTTTCAGCACAGAAATAAACA 467
Db	
QY	644 CCACCTCCGTTTATAGTAGGACCTCTGTTTCCAAATCTGSAATAACCCCTCACTCAAAAC 703
Db	
QY	468 CCACCTCTGTTTATAGTAGGACCTC--TTTCCAAATCTGSAATAACCCCTCACTCAAAAC 524
Db	
QY	704 TCACCTGTGTAAATTTAGCNAATCTATAGCACAAACAGCTCCCAATGCATCAGGTGG 763
Db	
QY	525 TCACCTGTGTAAATTTAGCNAATCTATAGCACAAACAGCTCCCAATGCATCAGGTGG 584
Db	
QY	764 TAACACCTCCACACGATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTTACCT 823
Db	
QY	585 TAACCTCTCCACACGATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTTACCT 644
Db	
QY	824 CAGCCTATCATTTGATGGCTCTTTCAGAACTATGTGCTCTCTCTCATTTCTTAGTGC 883
Db	
QY	645 CAGCCTATCATTTGATGGCTCTTTCAGAACTATGTGCTCTCTCTCATTTCTTAGTGG 704
Db	
QY	884 CCCCTATGACCATCTACCTGAAACAGATTTATACATCATGTCTACCTTAAGCCCCACA 943
Db	
QY	705 CCCCTATGACCATCTACCTGAAACAGATTTATACATCATGTCTACCTTAAGCCCCGCA 764
Db	
QY	944 ACAAGAGTAGTACCATCTCTCTTTGTTATCAGACGAGGAGTCTAGGACAGCTAGGTA 1003
Db	
QY	765 ACAAGAGTAGTACCATCTCTCTTTGTTATGAGACGAGGAGTCTAGGACAGCTAGGTA 824
Db	
QY	1004 CTGGCATTTGGCAGTATCACACCTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAA 1063
Db	
QY	825 CTGGCATTTGGGATATCACACCTCTACTCAGTTTCTACTCAAACTGTCTCAAGAAATAA 884
Db	
QY	1064 ATGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTTCGAAGATCACTTAACCTCCC 1123
Db	
QY	885 ATGTGACATGGAATGGTGGTCTGATACCTGGTCACTTTCGAAGATCACTTAACCTCCC 944
Db	
QY	1124 TAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAAGAGGG 1183
Db	
QY	945 TAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAAGAGGG 1004
Db	
QY	1184 GAACCTGTTTATTTTATAGGAGAAACCGCTGTTATTTATGTTAAATCAATPCCAGAAATGTC 1243
Db	
QY	1005 GAACCTTATTTTATAGGAGAAATGCTGTTGTTATGTTAAATCAATCCGAAATCATCA 1064
Db	

QY	1244 CTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGACGAGGAGCTTCAAAACA 1303
Db	
QY	1065 CCGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGACGAGGAGCTTCAAAACA 1124
Db	
QY	1304 CCGAACGCTGGGCGCTCTCCAGCAATGGAATGCGCTGGGTCTCCCTCTTCTTAGGACCTC 1363
Db	
QY	1125 CTGGACCTCTGGGCGCTCTCCAGCAATGGAATGCGCTGGGTCTCCCTCTTCTTAGGACCTC 1184
Db	
QY	1364 TAGCAGCTCTAATATATGTTACTCTCTTTGGACCTCTGTATCTTTAACTCTCTTTAAAGT 1423
Db	
QY	1185 TAGCAGCTAATATATGTTACTCTCTTTGGACCTCTGTATCTTTAACTCTCTTTAAAGT 1244
Db	
QY	1424 TTGTCTCTTCCAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1481
Db	
QY	1245 TTGTCTCTTCCAGAAATCGAAGCAGTAAAACTCAAAATCGTCTTCTCAAAATGGAGCCCCA 1302
Db	

RESULT 31

ABA56337	standard; DNA; 1894 BP.
ID	ABA56337
XX	
AC	ABA56337;
XX	
DT	01-FEB-2002 (first entry)
XX	
DE	Human foetal liver single exon nucleic acid probe #4642.
XX	
KW	Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
OS	Homo sapiens.
XX	
PN	W0200157277-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-05000669.
XX	
PR	04-FEB-2000; 2000US-0180312P.
PR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
XX	WPI; 2001-483447/52.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human foetal liver.
XX	
PS	Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
XX	
CC	The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention.
CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
'SQ	Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match	76.4%; Score 1131.6; DB 4; Length 1894;
Best Local Similarity	92.4%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;	
QY	164 TATCTAAGGAAACTCCACCTTCACTGCCACACCATATGCCCGCAACTGCTATAACT 223
Db	

|||||
1 TATCTAAGGAACCCACCTTCACTGCCACACCCATATGCCACAACTGCTATAACT 60
224 CTGCCACTCTTTGCAATGCAATCAATATCTATTAATGGACAGGAAATGATTAATCTTA 283
Db CTGCCACTCTTTGCAATGCAATCAATATCTATTAATGGACAGGAAATGATTAATCTTA 120
284 GTTGCTCTGGAGACTTGGAGCACTGCTGTGTGGACTTACTTCAACCCATACAGTATGT 343
Db GTTGCTCTGGAGACTTGGAGCACTGCTGTGTGGACTTACTTCAACCCATACAGTATGT 167
344 CTGATGGGGGTGGAATTCAGAGTCAGGCAAGAGAAACAACTAAAGGAAGCAATCTCC 403
Db CTGATGGGGGTGGAATTCAGAGTCAGGCAAGAGAAACAACTAAAGGAAGTAACTCC 227
404 AACTGACCCGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTTCCTCAAAACTAC 463
Db AACTGACCTGGGTACATAGCACCCCTGGCCCTACAAAGGACTAGATCTCTCAAAACTAC 287
464 ATGAACCTCTCGTACCCATCTCGCTGGTGGCTATTTAATACCAACCTCACTCGGC 523
Db ATGAACCTCTCGTACCCATCTCGCTGGTGGCTATTTAATACCAACCTCACTCGGC 347
524 TCATGAGGTCTCAGCCCAAAACCTTAACTGTGGATGTGCCTCCCTCGCACTTCA 583
Db TCATGAGGTCTCGGCCCAAAACCTTAACTGTGGATGTGCCTCCCTCGCACTTCA 407
584 GGCCATACATTTCAATCCCTGTCTGCTGAACAATGGAACTTTCAGCACAGAAATAACA 643
Db GGCCATACATTTCAATCCCTGTCTGCTGAACAATGGAACTTTCAGCACAGAAATAACA 467
644 CCACCTTCCTTTTAGTAGGCTCTGTTTCCAACTGGAAATACCCATACCTCAACC 703
Db CCACCTTCCTTTTAGTAGGCTCTGTTTCCAACTGGAAATACCCATACCTCAACC 524
704 TCACCTGTCTGAAATTTAGCAATCTATAGACACACCAAGCTCCCAATCATCAGTGGG 763
Db TCACCTGTCTGAAATTTAGCAATCTATAGACACACCAAGCTCCCAATCATCAGTGGG 584
764 TAACACCTCCCAACGAATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGGTACCT 823
Db TAACCTCCCTCCCAACGAATAGTCTGCTACCTCAGGAATATTTTTGTCTGTGGTACCT 644
824 CAGCCTATCATTTGTTGAATGGCTCTTCAGAACTATGTGCTTCTCATTTCTTAGTGC 883
Db CAGCCTATCATTTGTTGAATGGCTCTTCAGAACTATGTGCTTCTCATTTCTTAGTGC 704
884 CCCTATGACCATCTACACTGAACAAGATTTATACAAATCATGTCTGTAAGCCCAACA 943
Db CCCTATGACCATCTACACTGAACAAGATTTATACAAATCATGTCTAAGCCCAACA 764
944 ACAAAGAGTACCAATCTCTCTTTGTTATCAGACGAGAGTGTAGGACAGACTAGGTA 1003
Db ACAAAGAGTACCAATCTCTCTTTGTTATTTGGAGCAGGAGTGTAGGCGAGTAGCTA 824
1004 CTGGCAATTCGAGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAA 1063
Db CTGGCAATTCGAGTATCAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAA 884
1064 ATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAACTCC 1123
Db ATGGTGACATGGAATGGTGGTGGTATACCTGGTCACTTGGCAAGATCAACTTAACTCC 944
1124 TAGCAGAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGG 1183
Db TAGCAGAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCGGAAGAGGG 1004
1184 GRACCTGTTATTTTAGGAGAGAGCGCTGTTATTTATGTTAATCAATCAGAAATGTCA 1243
Db GRACCTGTTATTTTAGGAGAGAGATGCTGTTGTTATGTTAATCAATCGGAATCATCA 1064
1244 CTGAGAAAGTTAAGAAATTCGAGATCGAATCAATGTAGACAGAGGAGCTTCAAAACA 1303

Db 1065 CCGAGAAAGTTAAAGAAATTCAGGTGCAATATATAACGTAGACGAAGAGCTGCAAAACA 1124
Qy 1304 CCGAACGCTGGGGCTCTCTAGCCAAATGGATGCCCTGGTCTCCCTCTTTAGGACCTC 1363
Db 1125 CTGGACCTCTGGGGCTCTCTAGCCAAATGGATGCCCTGGATCTCCCTCTTTAGGACCTC 1184
Qy 1364 TAGCAGCTCTAATATTTGTTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTTAACT 1423
Db 1185 TAGCAGCTATATATTTGTTACTCTCTTTGGACCCCTGTATCTTTAACTCTCTTTAACT 1244
Qy 1424 TTGTCTCTTCAGAAATTCAGCTGTAAAGCTTACAGATGGTCTTACAAATGGAAACCCA 1481
Db 1245 TTGTCTTTTCCAGATCGAAGCAGTAAACTACAAATCGTTCTTCAAAATGGAGCCCA 1302

RESULT 32
AAI35980
ID AAI35980 standard; DNA; 1894 BP.
XX
AC AAI35980;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4666 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632386.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
(MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta.
PT
XX
PS Claim 25; SEQ ID NO 4666; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 76.4%; Score 1131.6; DB 4; Length 1894;
Best Local Similarity 92.4%; Pred. No. 0;
Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;

Qy 164 TATCTAAGGAACTCCACCTTCACTGGCCACACCCATATGCCCGCAACTGCTATAACT 223
Db 1 TATCTAAGGAACTCCACCTTCACTGGCCACACCCATATGCCCGCAACTGCTATAACT 60
Qy 224 CTGCCACTCTTTGCAATGCAATCAATATCTATTAATGGACAGGAAATGATTAATCTTA 283

Db 61 CTGCCACTCTTTGCATGTCATGCAAAATACATTAATTTGGACAGGAAAAAGATTAATCCCA 120
 Qy 284 GTTGTCTCTGAGACATTTGAGCCACTGCTGTGTTGGACTTACTTACCCATACCAATATGT 343
 Db 121 GTTGTCTGAGACATTTGGA-----GGACTCACTTCACTATACCAATATGT 167
 Qy 344 CTGATGGGGTGGAAATTTCAAGGTGAGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCC 403
 Db 168 CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATAAAGGAAGTAAATCTCC 227
 Qy 404 AACTGACCCGGGACATAGACACCCCTAGCCCTTACAAAGCACTAGTCTCTCAAACTAC 463
 Db 228 AACTGACCTGGGTACATAGACACCCCTGGCCCTTACAAAGCACTAGATCTCTCAAACTAC 287
 Qy 464 ATGAACCCCTCCGTACCCATCTGCGCTGTGAGCCCTATTATACCAACCTCACTCGGC 523
 Db 288 ATGAACCCCTCCATACCCATCTGCGCTGTGAGCCCTATTATACCAACCTCACTCGGC 347
 Qy 524 TCCATGAGGTCTCAGCCCAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCATCTCA 583
 Db 348 TCCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATGTGCTCCCTCCCTGCATCTCA 407
 Qy 584 GGCATACATTTCAATCCCTGTTCTGAAACAATGGAACAACTTCAGACACAGAAATTAACA 643
 Db 408 GGCATACATTTCAATCCCTATACCTGAAACAATGGAACAACTTCAGACACAGAAATTAACA 467
 Qy 644 CCATCTCCGTTTAGTAGGACCTCTGTTTCCATCTGGAATAAACCATACCTCAAAACC 703
 Db 468 CCATCTCTGTTTAGTAGGCTCTC---TTTCCATCTGGAATAAACCATACCTCAAAACC 524
 Qy 704 TCACCTGTGAAAAATTTAGCAATACTATAGACACAAACAGCTCCCAATGATCAGGTGG 763
 Db 525 TCACCTGTGAAAAATTTAGCAATACTATAGACACAAACAGCTCCCAATGATCAGGTGG 584
 Qy 764 TAACACCTCCACACGAATAGTGTGCTCACTCCCTCAGGAATATTTTTGTCTGTGGTACCT 823
 Db 585 TAACCTCTCCACACGAATAGTGTGCTCACTCCCTCAGGAATATTTTTGTCTGTGGTACCT 644
 Qy 824 CAGCTATCATTTGTTGAATGGCTCTTCAGAACTATGCTGCTCTCTCTATTTAGTGC 883
 Db 645 CAGCTATCATTTGTTGAATGGCTCTTCAGAACTATGCTGCTCTCTCTATTTAGTGC 704
 Qy 884 CCCCTATGACCATCTACACTGAAACAAGATTTATACAACTATGCTGCTACCTAAAGCCCA 943
 Db 705 CCCCTATGACCATCTACACTGAAACAAGATTTATACAACTATGCTGCTACCTAAAGCCCA 764
 Qy 944 ACAAAGAGTACCCATTTCTCTTTGTTATCAGACGAGAGTCTAGGAGACATAGGTA 1003
 Db 765 ACAAAGAGTACCCATTTCTCTTTGTTATTTGGAGCAGGAGTCTAGGCGGAGTAGCTA 824
 Qy 1004 CTGCAATGGGAGTATCAACACCTCTACTAGTTCTACTCAAACTATCTCAAGAAATAA 1063
 Db 825 CTGCAATGGGAGTATCAACACCTCTACTAGTTCTACTCAAACTATCTCAAGAAATAA 884
 Qy 1064 ATGTGATCATGGAACAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
 Db 885 ATGTGATCATGGAATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
 Qy 1124 TAGCAGCAGTATGCTCTTCAAAATCGAAGAGCTTTAGACTGCTTAACCGCAAAAGAGGG 1183
 Db 945 TAGCAGCAGTATGCTCTTCAAAATCGAAGAGCTTTAGACTGCTTAACCGCGGAAAGCGGG 1004
 Qy 1184 GAACCTGTTATTTTGGAGAGAAACGCTGTTATTTATGTTATTTATGTTATTTATGTTATTT 1243
 Db 1005 GAACCTGTTATTTTGGAGAGAAATGCTGTTGTTATTTATGTTATTTATGTTATTTATGTT 1064
 Qy 1244 CTGAGAAATTTAAAGAAATTCGAGATCGAATACAAATGAGAGCAGAGCTTCAAAACA 1303
 Db 1065 CCGAGAAATTTAAAGAAATTTCAAGTTCGAATATACAGTATGAGCAAGAGAGCTGCAAAACA 1124
 Qy 1304 CCGAACGCTGGGGCTCTCTAGCCAAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACCTC 1363
 Db 1125 CTGGACCTGGGGCTCTCTAGCCAAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACCTC 1184

Qy 1364 TAGCAGCTTAATATTTGTTACTCTCTTTGGACCCCTGATATCTTTAACTCTCTTTAAGT 1423
 Db 1185 TAGCAGCTTAATATTTGTTACTCTCTTTGGACCCCTGATATCTTTAACTCTCTTTAAGT 1244
 Qy 1424 TTGTCTCTTCCAGATTTGAAGCTGTAAAGCTACAGATGCTTTACAAATGGAACCCCA 1481
 Db 1245 TTGTCTCTTCCAGATTCGAAGCAGTAAACTCAAAATCGTTCTTCAATGGAGCCCA 1302

RESULT 33

AB45822 standard; DNA; 1894 BP.

XX ABA45822;

DT 01-FEB-2002 (first entry)

XX Human breast cell single exon nucleic acid probe #4517.

XX Human; microarray; single exon probe; gene expression; breast; disease;
 cancer; ss.

XX Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000662.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234587P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-496933/54.

XX New spatially-addressable set of single exon nucleic acid probes, useful
 for measuring gene expression in sample derived from human breast,
 comprises number of single exon nucleic acid probes.

XX Claim 1; SEQ ID NO 4517; 327pp + Sequence Listing; English.

XX The invention relates to a spatially-addressable set of single exon
 nucleic acid probes for measuring gene expression in a sample derived
 from human breast and BT 474 cells. The method involves contacting the
 probes with a collection of detectably labelled nucleic acids derived
 from mRNA of human breast, and then measuring the label bound to each
 probe of the microarray. The probes are useful for verifying the
 expression of regions of genomic DNA predicted to encode proteins. They
 are useful for gene discovery, and for determining predisposition and/or
 prognosing breast disease. Gene expression analysis is useful for
 assessing the toxicity of chemical agents on cells. The microarray of
 this invention presents a far greater diversity of probes for measuring
 gene expression, with far less bias than expressed sequence tag
 microarrays. The method is suitable for rapid production of functional
 information from genomic sequence. The present sequence is a single exon
 nucleic acid probe of the invention. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

XX Query Match

76.4%; Score 1131.6; DB 4; Length 1894;

Best Local Similarity 92.4%; Pred. No. 0;		Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;	
164	TATCTAAGGGAACCTCCACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACT	223	
1	TATCTAAGGGAACCCCACTTCACTGCCACACCCATATGCCCGCACTGCTATAACT	60	
224	CTGCCACCTCTTTGCATGCAATPACTCATTTATTTGGACAGGGAATAATTAATCTTA	283	
61	CTGCCACCTCTTTGCATGCAATPACTCATTTATTTGGACAGGGAATAATTAATCTTA	120	
284	GTTGTCTCGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCATACCAATATGT	343	
121	GTTGTCTCGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCATACCAATATGT	167	
344	CTGATGGGGGTGGAAATCAAGGTTCAGGCAAGAGAAAAAAGTAAAGGAAGCAATCTCCC	403	
169	CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACACATTAAGGAAGTAACTCCC	227	
404	AACTGACCCGGGACATAGCACCCTTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTAC	463	
228	AACTGACCTGGGTACATAGCACCCTTGCCCTTACAAAGGACTAGTCTCTCAAAACTAC	287	
464	ATGAAACCTTCGTATCCCATCTCGCTGGTGGAGCCTATTTAATACCAACCTCAGTCGGC	523	
288	ATGAAACCTTCATACCCCATCTGGCTGGTGAAGCCTATTTAATACCAACCTCAGTCGGC	347	
524	TCCATGAGGTCTCAGCCCAAAACCCCTACTTAACCTGTTGGATGTCCTCCCTCGACTTCA	583	
348	TCATGAGGTCTCGGCCCAAAACCCCTACTTAACCTGTTGGATGTCCTCCCTCGACTTCA	407	
584	GGCCATACATTTCAATCCCTGTTTCTTGAAACAATGGAACTTTCAGCACAGAAATAACA	643	
408	GGCCATACATTTCAATCCCTTATACCTGAACAATGGAACTTTCAGCACAGAAATAACA	467	
644	CAACTTCCTGTTTAGTAGGACTCTTGTCTTTCCTGTTTCCAACTCTGGAATAAACCCTA	703	
468	CAACTTCCTGTTTAGTAGGACTCTTGTCTTTCCTGTTTCCAACTCTGGAATAAACCCTA	524	
704	TCACCTGTCTTAAATTTAGCAATACCTATAGACAAACCACTCCCAATGATCAGGTGGG	763	
525	TCACCTGTGTAAATTTAGCAATACCTATAGACAAACCACTCCCAATGATCAGGTGGG	584	
764	TAAACCTCCCAACGAATAGTCTGCTACCTCCTCAGGAATATTTTGTCTGTGTACCT	823	
585	TAACTCTCCCAACGAATAGTCTGCTACCTCCTCAGGAATATTTTGTCTGTGTACCT	644	
824	CAGCCTATCATTTGTTGAATGGCTCTTTCAGAACTCTATGTGCTTCTCTCATTTCTAGTC	883	
645	CAGCCTATCATTTGTTGAATGGCTCTTTCAGAACTCTGTGTGCTTCTCTCATTTCTAGTC	704	
884	CCCTATGACCATCTACCTGAACAGATTTTATCAATCATGTCTGTAACCTTAAGCCCCACA	943	
705	CCCTATGACCATCTACCTGAACAGATTTTATCAATCATGTCTGTAACCTTAAGCCCCACA	764	
944	ACAAAAGAGTACCAATTTCTTCTTTTGTATCAGACGAGGAGTCTAGGCAGACTAGGTA	1003	
765	ACAAAAGAGTACCAATTTCTTCTTTTGTATTTGGACGAGGAGTCTAGGCAGACTAGGTA	824	
1004	CTGGCATGGCAGTATCAAACTCTACTCAGTTCTACTAATACTATCTCAAGAAATAA	1063	
825	CTGGCATGGCGGTATCAAACTCTACTCAGTTCTACTAATACTATCTCAAGAAATAA	884	
1064	ATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGGCAGATCAACTTAACCTCC	1123	
885	ATGGTGACATGGAATGGGTCGCTGATACCTGGTCACTTGGCAGATCAACTTAACCTCC	944	
1124	TAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTTGCTTAAACCGCAAGAGGGG	1183	
945	TAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTTGCTTAAACCGCGAAGCGGG	1004	
1184	GAACTCTGTTTATTTTAGGAGAAAGACGCTGTTATTTATGTTAATCAATTCAGAAATGCTCA	1243	

Qy	164	TATCTAAGGGAAACTCCAACCTTCTACTGCCACACCCATATGCCCGCAACTGCTATAACT	223
Db	1	TATCTAAGGGAAACCCCCACCTTCACTGCCACACCCCATATGCCCCCAACTGCTATAACT	60
Qy	224	CTGCCACTTTTGCATGCATGCAAAATACTCATTTATTGGACAGGGAATAATGATTAATCCTTA	283
Db	61	CTGCCACTTTTGCATGCATGCAAAATACTCATTTATTGGACAGGGAATAATGATTAATCCTTA	120
Qy	284	GTTGTCTCTGAGGACTTTGGAGCCACTGTCTGTTGGACTTACTTCACCCATACACAGATATGT	343
Db	121	GTTGTCTCTGAGGACTTTGGA-----GGACTCACTTCACTCATACACAGATATGT	167
Qy	344	CTGATGGGGTGGAAATTCAGAGTCAAGCAAGAGAGAAAAACAAGTAAGAGCAATCTCCC	403
Db	168	CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAAAACAATAAAGGAAGTAATCTCCC	227
Qy	404	AATGACCCGGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTAC	463
Db	228	AACTGACTGGGTACATAGCACCCCTGGCCCTACAAAGGACTAGATCTCTCAAAACTAC	287
Qy	464	ATGAAACCCCTCCGTAACCCACTACTCGCTGTGTAGCCCTATTTAATACCAACCTCACTCGGC	523
Db	288	ATGAAACCCCTCCATACCCCATACTGGCTGTGTAGCCCTATTTAATACCAACCTGACTGGGC	347
Qy	524	TCCATGAGGTCTCAGCCGCAAAACCTACTAACTGTTGGATGTGCTCTCCCTCGCACTTCA	583
Db	348	TCCATGAGGTCTCGGCCCAAAACCTACTAACTGTTGGATGTGCTCTCCCTCGCACTTCA	407
Qy	584	GGCCATACATTTCAATCCCTGTTCTGAAACATGGAACAACCTTCAGCACAGAAATAACA	643
Db	408	GGCCATACATTTCAATCCCTATACCTGAACAATGGAAACAACCTTCAGCACAGAAATAACA	467
Qy	644	CCACTTCGGTTTTAGTAGGACCTCTGTTTCCAACTCTGGAAATAACCCATCACTCAAAAC	703
Db	468	CCACTTCGTTTTAGTAGGTCTCTC--TTTCCAACTCTGGAAATAACCCATCACTCAAAAC	524
Qy	704	TCACCTGTGTAAAAATTTAGCAATACTATAGACACAACACAGCTCCCAATGANTCAGGTGGG	763
Db	525	TCACCTGTGTAAAAATTTAGCAATACTATAGACACAGCCAACCTCCCAATGANTCAGGTGGG	584
Qy	764	TAAACACTCCACACAGAAATAGTCGCTACCTCAGGAATATTTTTGTCTGTGCTACCT	823
Db	585	TAACTCTCTCCACACAGAAATAGTCGCTACCTCAGGAATATTTTTGTCTGTGCTACCT	644
Qy	824	CAGCCTATCATTTGTAATGGCTCTTCAGAAATCTATGTGCTCTCCTCTCATTTCTTAGTGC	883
Db	645	CAGCCTATCATTTGTAATGGCTCTTCAGAAATCTATGTGCTCTCCTCTCATTTCTTAGTGC	704
Qy	884	CCCCTATGACCATCTACACTGAAACAAGATTTATACATCATGTGCTACCTTAAGCCCCACA	943
Db	705	CCCCTATGCCCCATCTACACTGAAACAAGATTTATCAATCATGTCTACCTTAAGCCCCGCA	764
Qy	944	ACAAAGAGTACCCATCTCTCTTTTGTATCAGACAGGAGTCTAGGCAGACTAGGTA	1003
Db	765	ACAAAGAGTACCCATCTCTCTTTTGTATTTGGACAGGAGTCTAGGCAGGAGTACTA	824
Qy	1004	CTGGCATTTGGCAGTATCACAACTCTACTCAGTTTCTACTCAAACTATCTCAAGAAATAA	1063
Db	825	CTGGCATTTGGCAGTATCACAACTCTACTCAGTTTCTACTCAAACTGTCTCAAGAAATAA	884
Qy	1064	ATGCTGACATGGAAACAGGTCACTGACTCCCTGGTCACTTTGGCAGAGTCMACTTAACCTCC	1123
Db	885	ATGCTGACATGGAAATGGGTGCGCTGATACCTCTGGTCACTTTGGCAGAGTCAACTTAACCTCC	944
Qy	1124	TAGCAGCAGTATGCTCTTCAAAATCGAAGAGCTTTTAGACTTCTTAACCGCCCAAAAGAGGGG	1183
Db	945	TAGCAGCAGTATGCTCTTCAAAATCGAAGAGCTTTTAGACTTCTTAACCGCGAAAGCGGG	1004
Qy	1184	GAACTGTTTATTTTATGGAGAGAAACGCTGTTATTTATGTTTAATCAATCCAGAAATGTCA	1243
Db	1005	GAACTGTTTATTTTATGGAGAAATGCTGTGTTATTTATGTTTAATCAATCCGGAATCATCA	1064

Qy	1244	CTGAGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACA	1301			
Db	1065	CCGAGAAAGTTAAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACA	1124			
Qy	1304	CCGAAACGCTGGGGCCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCTCTTTAGGACCTC	1363			
Db	1125	CTGACACCTGGGGCCCTCCTCAGCCAATGGATGCCCTGGGTTCTCCCTCTTTAGGACCTC	1184			
Qy	1364	TAGCAGCTCTAAATATTTGTTACTCTCTTTTGGACCCCTGTATCTTTTAACCTCTCTTTAAAGT	1423			
Db	1185	TAGCAGCTCTAAATATTTGTTACTCTCTTTTGGACCCCTGTATCTTTTAACCTCTCTTTAAAGT	1244			
Qy	1424	TTGTCTCTTCCAGAAATGAAGCTGTAAAGCTCAGATGGTCTTACAAATGGAAACCCCA	1481			
Db	1245	TTGTCTTTCAGAAATGAAGCTGTAAAGCTCAGATGGTCTTACAAATGGAAACCCCA	1302			
RESULT 36						
ID	AAK04516 standard; DNA; 1894 BP.					
XX	AAK04516;					
XX	AAK04516;					
DT	05-NOV-2001 (first entry)					
DE	Human brain expressed single exon probe SEQ ID NO: 4507.					
KW	Human; brain expressed exon; gene expression analysis; probe; microarray;					
KW	Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;					
XX	ss.					
OS	Homo sapiens.					
XX	WO200157275-A2.					
XX	09-AUG-2001.					
XX	30-JAN-2001; 2001WO-US000667.					
XX	04-FEB-2000; 2000US-0180312P.					
XX	26-MAY-2000; 2000US-0207456P.					
XX	30-JUN-2000; 2000US-00608408.					
XX	03-AUG-2000; 2000US-00632366.					
XX	21-SEP-2000; 2000US-0234687P.					
XX	27-SEP-2000; 2000US-0236359P.					
XX	04-OCT-2000; 2000GB-00024263.					
XX	(MOLE-) MOLECULAR DYNAMICS INC.					
XX	Penn SG, Hanzel DK, Chen W, Rank DR;					
XX	WPI; 2001-483446/52.					
XX	Single exon nucleic acid probes for analyzing gene expression in human					
XX	brains.					
XX	Example 4; SEQ ID NO 4507; 650pp + Sequence Listing; English.					
XX	The present invention provides a number of single exon nucleic acid					
XX	probes which are derived from genomic sequences expressed in the human					
XX	brain. They can be used to measure gene expression in brain cell samples					
XX	which may enable the diagnosis and improved treatment of nervous system					
XX	diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,					
XX	epilepsy and cancers. The present sequence is one of the probes of the					
XX	invention					
XX	Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;					
Query Match 76.4%; Score 1131.6; DB 4; Length 1894;						
Best Local Similarity 92.4%; Pred. No. 0;						
Matches 1218; Conservative 0; Mismatches 84; Indels 16; Gaps 2;						
Oy	164 TATCTAGGAAATCTCCACCTTCTACGCCACCACTATGCCCGCAACTGTCTATAACT					
223						

Query Match		76.4%;	Score 1131.6;	DB 4;	Length 1894;	
Best Local Similarity		92.4%;	Pred. No. 0;			
Matches 1218;		Conservative	0;	Mismatches	84;	Indels 16; Gaps 2;
QY	164	TATCTAAGGAAACTCA	CTTCACTGCGCCACACCCATATGCCCCCGCAACTGCTATACT	223		
Db	1	TATCTAAGGAAACCC	CACTTCACTGCGCCACACCCATATGCCCCCACTGCTATACT	60		
QY	224	CTGCCACTCTTTCAT	GCATGCAATCACTTATTTGGACGGGAAATGATTAATCCTA	283		
Db	61	CTGCCACTCTTTCAT	GCATGCAATCACTTATTTGGACGGGAAATGATTAATCCTCA	120		
QY	284	GTGTGCTCTGGAGAC	TGAGGCCACTGTCTGTGTGGACTTACTTTCACCCATACAGTATGT	343		
Db	121	GTGTGCTCTGGAGAC	TTGGA-----GGACTCACCTTCACTCATACAGTATGT	167		
QY	344	CTGATGGGGTGAAAT	TCAAGGTTCAGGCAAGAGAAACAAAGTAAAGGAAGCAATCTCCC	403		
Db	168	CTGATGGGGTGAGAT	TCAAGATCAGGCAACAGAAACACATATAAGGAAGTAAATCTCCC	227		
QY	404	AACCTGACCCGGGAC	ATAGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTAC	463		
Db	228	AACCTGACCCGGGAC	ATAGCACCCCTAGCCCTACAAAGGACTAGTATCTCAAAACTAC	287		
QY	464	ATGAAACCCCTCGT	ATACCCACTGCGCTGGTGAGCCCTATTTAATACACCCCTCACTCGGC	523		
Db	288	ATGAAACCCCTCGT	ATACCCACTGCGCTGGTGAGCCCTATTTAATACACCCCTCACTCGGC	347		
QY	524	TCCATGAGTCTCAG	CCCAAAACCCCTACTAACTGTGTGGATGTGCTCCCTCGCTGCTTCA	583		
Db	348	TCCATGAGTCTCAG	CCCAAAACCCCTACTAACTGTGTGGATGTGCTCCCTCGCTGCTTCA	407		
QY	584	GGCCATACATTTCA	ATCCCTCTCTGGAACATGGAACAACTTCAGCACAGAAATAAACA	643		
Db	408	GGCCATACATTTCA	ATCCCTCTCTGGAACATGGAACAACTTCAGCACAGAAATAAACA	467		
QY	644	CCACTTCGCTTTAG	TAGGACCTCTTGTTCCTCAATCTGGAAATACCCATACCTCAAAAC	703		
Db	468	CCACTTCGCTTTAG	TAGGACCTCTTGTTCCTCAATCTGGAAATACCCATACCTCAAAAC	524		
QY	704	TCACCTGTGTAAAT	TTAGCAATACCTATAGACACAAACAGCTCCCAATCATCAGTGGG	763		
Db	525	TCACCTGTGTAAAT	TTAGCAATACCTATAGACACAAACAGCTCCCAATCATCAGTGGG	584		
QY	764	TAAACCTCCCAACA	AGATGTGCTACCTCAGGAAATATTTTGTGCTGTGCTACCT	823		
Db	585	TAACTCTCCCAACA	AGATGTGCTACCTCAGGAAATATTTTGTGCTGTGCTACCT	644		
QY	824	CAGCCTATCATGTT	GAATGGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTTAGTGC	883		
Db	645	CAGCCTATCATGTT	GAATGGCTCTTCAGAAATCTATGTGCTTCCTCTCATTTCTTAGTGC	704		
QY	884	CCCTATGACCATCT	ACATGACAGATTTATACATCATCTGCTACCTTAAGCCCAACA	943		
Db	705	CCCTATGACCATCT	ACATGACAGATTTATACATCATCTGCTACCTTAAGCCCAACA	764		
QY	944	ACAAAGAGTACCC	ATCTCTCTTTGTTATCAGACGAGGTGCTAGGACAGTACCTAGGTA	1003		
Db	765	ACAAAGAGTACCC	ATCTCTCTTTGTTATGAGGAGGAGTGTAGGCGGAGTAGCTA	824		
QY	1004	CTGGCAATGGCAGT	ATCAAACTCTACGTTCTACTACAAACTATCTCAAGAAATAA	1063		
Db	825	CTGGCAATGGCAGT	ATCAAACTCTACGTTCTACTACAAACTGTCTCAAGAAATAA	884		
QY	1064	ATGGTGACATGNA	CAGGTCACTGACTCCCTGGTCACTTGCAGAGTCAACTTAATCCC	1123		
Db	885	ATGGTGACATGNA	CAGGTCACTGACTCCCTGGTCACTTGCAGAGTCAACTTAATCCC	944		
QY	1124	TAGCAGCAGTAGT	CTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAGAGGGG	1183		
Db	945	TAGCAGCAGTAGT	CTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGGGAAGCGGG	1004		
QY	1184	GAACTGTATTATTT	TAGGAGAGAAACGCTGTATTATTTATGTTAATCAATCCAGAAATGTCA	1243		

Db	1005	GAACCTTTTATTTT	TAGAGGAAAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCA	1064
Qy	1244	CTGAGAAAGTTAA	GAAATTCGAGATCGAATACAAATGTAGACGAGAGAGCTTCAAAACA	1303
Db	1065	CCGAGAAAGTTAA	GAAATTCAGAGTCGATATATACGTAGACGCAAGGAGCTGCAAAACA	1124
Qy	1304	CCGAACTCTGGG	CCCTCTCAGCAATGGATGCGCTGGGTTCTCCCTTCTTAGGACCTC	1363
Db	1125	CTGACCTCTGGG	CCCTCTCAGCAATGGATGCGCTGGATCTCCCTTCTTAGGACCTC	1184
Qy	1364	TAGCAGCTCTAA	TATGTTACTCCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGT	1423
Db	1185	TAGCAGCTATA	TATGTTACTCCTCTTTGGACCCCTGTATCTTTAACTCTCTTGTAAAGT	1244
Qy	1424	TTGTCTCTTCC	GAAATTCGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAAACCCCA	1481
Db	1245	TTGTCTTTC	AGAACTGAAGCACTAAACCTACAAATCGTCTTCAAAATGGAGCCCCA	1302
RESULT 38				
AAI04422				
ID	AAI04422 standard; DNA; 1894 BP.			
XX	AAI04422;			
XX	AC			
DT	09-OCT-2001 (first entry)			
XX	Probe #4413 used to measure gene expression in human breast sample.			
DE	Probe; human; breast disease; breast cancer; development disorder; es;			
KW	inflammatory disease; proliferative breast disease; non-carcinoma tumour.			
KW	Homo sapiens.			
OS				
XX	WO200157270-A2.			
XX	09-AUG-2001.			
XX	29-JAN-2001; 2001WO-US000661.			
XX				
PR	04-FEB-2000; 2000US-0180312P.			
PR	26-MAY-2000; 2000US-0207456P.			
PR	30-JUN-2000; 2000US-0060840P.			
PR	03-AUG-2000; 2000US-0063236P.			
PR	21-SEP-2000; 2000US-0234687P.			
PR	27-SEP-2000; 2000US-0236359P.			
PR	04-OCT-2000; 2000GB-00024263.			
XX	(MOLE-) MOLECULAR DYNAMICS INC.			
XX				
PI	Penn SG, Hanzel DK, Chen W, Rank DR;			
XX				
XX	WPI; 2001-476286/51.			
XX				
XX	Novel single exon nucleic acid probe used to measuring gene expression in			
PT	a human breast.			
PS	Claim 25; SEQ ID NO 4413; 322pp; English.			
XX				
XX	The present invention relates to novel single exon nucleic acid probes.			
CC	The present sequence is one such probe. The probes are useful for			
CC	measuring human gene expression in a human breast sample, where the probe			
CC	hybridises at high stringency to a nucleic acid expressed in the human			
CC	breast. The probes are useful for predicting, diagnosing, grading,			
CC	staging, monitoring and prognosing diseases of the human breast,			
CC	particularly those diseases with polygenic aetiology. The diseases			
CC	include: breast cancer, disorders of development, inflammatory diseases			
CC	of the breast, fibrocystic changes, proliferative breast disease and non-			
CC	carcinoma tumours. Note: The sequence data for this patent did not form			
CC	part of the printed specification, but was obtained in electronic format			
CC	directly from WIPO at ftp://wipo.int/pub/published_pct_sequences			
XX				

SQ	Sequence	1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match	76.4%;	Score 1131.6; DB 5; Length 1894;
Best Local Similarity	92.4%;	Pred. No. 0;
Matches 1218; Conservative	0;	Mismatches 84; Indels 16; Gaps 2;
QY	164	TATCTAAGGAAACTCCACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAA
DB	1	TATCTAAGGAACTCCACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAA
QY	224	CTGCCACTTTTGATGCAATGCAAACTCACTATTTGGACAGGAAATGATTAATCTTA
DB	61	CTGCCACTTTTGATGCAATGCAAACTCACTATTTGGACAGGAAATGATTAATCTTA
QY	284	GTGTCTCTGAGACACTGGAGCCACTGTCTGTGGACTTACTTCAACCATACAGATATGT
DB	121	GTGTCTCTGAGACACTGGG-----GGACTCACTTCACTATACAGATATGT
QY	344	CTGATGGGGTGGAAATTCAGAGGTGAGCAAGAGAAACAAAGTAAAGGAAGCAATCTCC
DB	168	CTGATGGGGTGGAGTTCAAGATCAGGCAACAGAAACACATAAAGGAAGTAACTCTCC
QY	404	AACGTACCCGGGACATAGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAAACTAC
DB	228	AACGTACCTGGGTACATAGCACCCCTGGCCCTACAAAGGACTAGTCTCTCAAACTAC
QY	464	ATGAAACCCCTCCGTACCCACTACTCGCTGGTGGCCCTATTATATACACCCCTCACTGGC
DB	288	ATGAAACCCCTCCATACCCACTACTGGCTGGTGGCCCTATTATATACACCCCTCACTGGC
QY	524	TCCATGAGGTCTCAGCCCAAAACCCCTACTAATCTGTGGATGTGCTCCCTCTGCACCTCA
DB	348	TCCATGAGGTCTCGGCCCAAAACCCCTACTAATCTGTGGATGTGCTCCCTCTGCACCTTA
QY	584	GGCCATACATTTCAATCCCTGTTCTCGAAACAATGGAACAATTCAGACAGAAATAACA
DB	408	GGCCATACATTTCAATCCCTTATACCTGAAACAATGGAACAATTCAGACAGAAATAACA
QY	644	CACTTCGGTTTGTAGTAGGACCTCTGTTCCCAATCTGGAATTAACCATACCTCAACCC
DB	468	CACTTCGGTTTGTAGTAGGACCTCTGTTCCCAATCTGGAATTAACCATACCTCAACCC
QY	704	TCACCTGTGTAATAATTTAGCAATATATAGACACAACAGCTCCCAATGCATCAGGTGGG
DB	525	TCACCTGTGTAATAATTTAGCAATATATAGACACAACAGCTCCCAATGCATCAGGTGGG
QY	764	TAAACCTCCACACGAAATAGTCTGCCTACCTCAGGAATATTTTTGTCTGTGTACCT
DB	585	TAACTCTCCACACGAAATAGTCTGCCTACCTCAGGAATATTTTTGTCTGTGTACCT
QY	824	CAGCTATCATTTGTAATGGCTCTTCAGNAATCATGTGCTTCTCTCATTTCTTAGTGC
DB	645	CAGCTATCATTTGTAATGGCTCTTCAGNAATCATGTGCTTCTCTCATTTCTTAGTGC
QY	884	CCCTATGACCATCTACACTGAAACAAGATTTATCAATCATGTCTACCTTAAGCCCAACA
DB	705	CCCTATGACCATCTACACTGAAACAAGATTTATCAATCATGTCTACCTTAAGCCCAACA
QY	944	ACAAAGAGTACCCATCTTCTTTGTTATCAGACGAGGAGTCTAGGACAGTAGGTA
DB	765	ACAAAGAGTACCCATCTTCTTTGTTATTTGGACGAGGAGTCTAGGCGGAGTAGGTA
QY	1004	CTGGCATGGGAGTATCAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAA
DB	825	CTGGCATGGGAGTATCAACCTCTACTAGTTCTACTACAACTATCTCAAGAAATAA
QY	1064	ATGGTGACATGGAACAGCTACTGCTCCCTGGTCACTTGCCTCAAGATCAACTTAACTCCC
DB	895	ATGGTGACATGGAATGGGTGCTGATACCTGGTCACTTGCCTCAAGATCAACTTAACTCCC
QY	1124	TAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGAACCGCCAAAGAGGGG
DB	945	TAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGAACCGCCGGAAGCGGGG

1184

GAACCTGTTTATTTTAGGAGAAACGCTGTTATTTATGTTATTAATCAATCCAGATTGTCA

1243

1005

GAACCTTTTTTATTTTAGGAGGAAAAATGCTGTTGTTATGTTATTAATCAATCCGGAATCATCA

1064

1244

CTGAGAAAGTTAAAGAAATTCAGATCGAATCAATGTTAGACGAGGAGCTTCAAACA

1303

1065

CCGAGAAAGTTAAAGAAATTCAGAGTTCGAATATAAGCTAGAGCAAGAGGAGCTGCAANACA

1124

1304

CCGAAAGCTGGGCGCTCTCCAGCAATGGATGATGCGCTGGGTTCTCCCTCTTTAGGACCTC

1363

1125

CTGGACCTTGGGCGCTCTCAGCCATGATGATGCGCTGGATCTCCCTCTTTAGGACCTC

1184

1364

TAGCAGCTCAATATTTGTTACTCTCTTTGGACCCCTGATCTTTAACTCTCTTTAAAGT

1423

1185

TAGCAGCTAATAATTTGTTACTCTCTTTGGACCCCTGATCTTTAACTCTCTTTAAAGT

1244

1424

TTGCTCTTCCAGAAATTTGAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAAACCCA

1481

1245

TTGCTCTTCCAGAAATTCGAAGCAGTAAACTAAAATCGTTCTTCAATGGAGCCCA

1302

RESULT 39

ABS04589

ID

ABS04589 standard; DNA; 1894 BP.

XX

ABS04589;

AC

19-AUG-2002 (first entry)

DT

Human genome-derived single exon probe from lung SEQ ID No 4580.

DE

Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX

chronic obstructive pulmonary disease; interstitial lung disease;

XX

familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX

tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;

XX

Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;

XX

pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX

pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX

primary ciliary dyskinesia; pulmonary hypertension;

XX

hyaline membrane disease.

OS

Homo sapiens.

XX

W0200186003-A2.

PN

15-NOV-2001.

PD

30-JAN-2001; 2001W0-US000665.

XX

04-FEB-2000; 2000US-0180312P.

XX

26-MAY-2000; 2000US-0207456P.

XX

30-JUN-2000; 2000US-00608408.

XX

03-AUG-2000; 2000US-00632366.

XX

21-SEP-2000; 2000US-0234687P.

XX

27-SEP-2000; 2000US-0236359P.

XX

04-OCT-2000; 2000GB-00024263.

XX

(MOLE-) MOLECULAR DYNAMICS INC.

PA

Penn SG, Hanzel DK, Chen W, Rank DR;

XX

WPI; 2002-114183/15.

XX

Spatially-addressable set of single exon nucleic acid probes, used to

XX

measure gene expression in human lung samples.

PT

Claim 1; SEQ ID NO 4580; 634pp; English.

XX

The invention relates to a spatially-addressable set of single exon

XX

nucleic acid probes for measuring gene expression in a sample derived

XX

from human lung comprising single exon nucleic acid probes having one of

XX

12614 nucleic acid sequences mentioned in the specification, or their

complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match	76.4%;	Score 1131.6;	DB 6;	Length 1894;
Best Local Similarity	92.4%;	Pred. No. 0;		
Matches 1218;	Conservative 0;	Mismatches 84;	Indels 16;	Gaps 2;
QY	164	TATCTAAGGGAACTCCACCTTCACTGCGCCACACCCATATGCGCGCAACTCTCTATACT	223	
DB	1	TATCTAAGGGAACTCCACCTTCACTGCGCCACACCCATATGCGCGCAACTCTCTATACT	60	
QY	224	CTGCGACTCTTTGCGATGCGAATACTCATTTTGGACGGGAAAATGATTATCTTA	283	
DB	61	CTGCGACTCTTTGCGATGCGAATACTCATTTTGGACGGGAAAATGATTATCTTA	120	
QY	284	GTGTGCTCGAGGACTTGGAGCCACTGTCTGTTGGACTTACTTTCACCCATACCAGTATGT	343	
DB	121	GTGTGCTCGAGGACTTGGAGCCACTGTCTGTTGGACTTACTTTCACCCATACCAGTATGT	167	
QY	344	CTGATGGGGGTGGAAATTCAGGTTCAGGCAAGAGAAAACAAGTAAAGGAAGCAATCTCCC	403	
DB	168	CTGATGGGGGTGGAGTTCAAGATCAGGCAACAGAAAACACATAAAGGAAGTAAATCTCCC	227	
QY	404	AACTGACCCGGGACATAGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTAC	463	
DB	228	AACTGACCTTGGGTACATAGCACCCCTGGGCCCTTACAAAGGACTAGATCTCTCAAAACTAC	287	
QY	464	ATGAAACCTCGTACCCCATCTCGCTGCTGAGGCTATTTAATACCAACCTCTACCTCGGC	523	
DB	288	ATGAAACCTCGTACCCCATCTCGCTGCTGAGGCTATTTAATACCAACCTCTACCTCGGC	347	
QY	524	TCCATGAGGTCTCAGCCCAAAACCTTACTAATCTGTTGGATGTCCTCCCTCGCACTTCA	583	
DB	348	TCCATGAGGTCTCAGCCCAAAACCTTACTAATCTGTTGGATGTCCTCCCTCGCACTTCA	407	
QY	584	GSCCATACATTTCAATCCCTGTTCTGTAACATGGAACACTTTCAGCACAGAAAATAACA	643	
DB	408	GSCCATACATTTCAATCCCTGTTCTGTAACATGGAACACTTTCAGCACAGAAAATAACA	467	

QY	644	CCACTTCGTTTGTAGTAGGACCTCTTGTTCCTCAATCTCGAAATACCCATACCTCAAAACC	703	
DB	468	CCACTTCGTTTGTAGTAGGTCCTC---TTTCCAACTCGAAATACCCATACCTCAAAACC	524	
QY	704	TCACCTGTGTAAATTTAGCAATATCTATAGACACAAACAGCTCCCAATGCAATCAGGTGGG	763	
DB	525	TCACTGTGTAAATTTAGCAATATCTATAGACACAGCAACTCCCAATGCAATCAGGTGGG	584	
QY	764	TAAACCTCCACACAGCAATAGTCTGCCTACCTCAGGAATATTTTCTCTGTGTACTCT	823	
DB	585	TAACTCTCCACACAGCAATAGTCTGCCTACCTCAGGAATATTTTCTCTGTGTACTCT	644	
QY	824	CAGCTATCATTTGTTGAATGGCTCTTCAGAAATCTATGTCTCTCTCATTTCTTAGTGC	883	
DB	645	CAGCTATCATTTGTTGAATGGCTCTTCAGAAATCTATGTCTCTCTCATTTCTTAGTGC	704	
QY	884	CCCTATGACCATCTACACTGAAACAGATTTATACAATCATGTCTGCTGCTTAAGCCCAACA	943	
DB	705	CCCTATGACCATCTACACTGAAACAGATTTATACAATCATGTCTGCTTAAGCCCAACA	764	
QY	944	ACAAAGAGTACCCATTTCTCTCTTTGTTATCAGAGCAGGAGTCTAGGCGAGACTAGTGA	1003	
DB	765	ACAAAGAGTACCCATTTCTCTCTTTGTTATGAGAGCAGGAGTCTAGGCGAGTACTGA	824	
QY	1004	CTGGCAATGGCAGTATCACAACTCTACTCAGTTCTACTACAAACTCTCTCAAGAAATAA	1063	
DB	825	CTGGCAATGGCAGTATCACAACTCTACTCAGTTCTACTACAAACTCTCTCAAGAAATAA	884	
QY	1064	ATGTTGACATGGACAGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1123	
DB	885	ATGTTGACATGGACAGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	944	
QY	1124	TAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGG	1183	
DB	945	TAGCAGCAGTACTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGG	1004	
QY	1184	GAACCTGTTTATTTTAGGAGAAACGCTGTATTATTTATTTATTTATTTATTTATTTAT	1243	
DB	1005	GAACCTGTTTATTTTAGGAGAAACGCTGTATTATTTATTTATTTATTTATTTATTTAT	1064	
QY	1244	CTGAGAAAGTTTAAAGAAATTCGAGATCGAATACAAATGATAGAGAGAGCTTCAAAACA	1303	
DB	1065	CCGAGAAAGTTTAAAGAAATTCGAGATCGAATACAAATGATAGAGAGAGCTTCAAAACA	1124	
QY	1304	CCGAGAGCTGGGGCTCTCTCAGCAATGGAATGCGCTCTCTCTCTCTCTCTCTCTCTCT	1363	
DB	1125	CTGAGAGCTGGGGCTCTCTCAGCAATGGAATGCGCTCTCTCTCTCTCTCTCTCTCTCT	1184	
QY	1364	TAGCAGCTCTAATATTTGTTACTCTCTCTTTGGACCTCTCTCTCTCTCTCTCTCTCTCT	1423	
DB	1185	TAGCAGCTCTAATATTTGTTACTCTCTCTTTGGACCTCTCTCTCTCTCTCTCTCTCTCT	1244	
QY	1424	TGTCTCTTCCAGAAATGGAAGCTGTAAGCTTACAGATGGTCTTCAAAATGGAACCCCA	1481	
DB	1245	TGTCTCTTCCAGAAATGGAAGCTGTAAGCTTACAGATGGTCTTCAAAATGGAACCCCA	1302	
RESULT 40				
AAAX25659				
ID	AAAX25659 standard; cDNA to mRNA; 1948 BP.			
XX				
AC	AAAX25659;			
XX				
DT	21-MAY-1999 (first entry)			
XX				
DE	Human endogenous retrovirus W clone cl.24.4.			
XX				
KW	Clone; human endogenous retrovirus; genome; autoimmune disease;			
KW	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;			
XX	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.			
OS	Human endogenous retrovirus.			
XX				

PN WO9902696-A1.
 XX 21-JAN-1999.
 XX 06-JUL-1998; 98WO-FR001442.
 XX 07-JUL-1997; 97FR-00008815.
 XX (INMR) BIO MERIEUX.
 XX Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
 XX WPI; 1999-120897/10.
 XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
 PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
 PT disease, and abnormal or failed pregnancy.
 XX Claim 1; Page 56-58; 106pp; French.
 XX This sequence represents clone cl.24.4 of the human endogenous retrovirus
 CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
 CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
 CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
 CC dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility
 XX SQ Sequence 1948 BP; 505 A; 549 C; 381 G; 510 T; 0 U; 3 Other;
 Query Match 75.4%; Score 1116.4; DB 2; Length 1948;
 Best Local Similarity 94.2%; Pred. No. 0;
 Matches 1159; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTATCATCTTTCTCTTACTGTTCTTCTTACCCCTTTTGGTCTCACT 60
 DB 719 ATGGCCCTCCCTATCATATTTTCTTACTGTTGTTTTCACCTCTTTCACCTCACT 778
 QY 61 GCACCCCTCGTCTGCTGTACACACAGTAGTCTCCCTTACCAGAGTTTCTATGAGA 120
 DB 779 GCACCCCTCGTCTGCTGTATGACAGTAGTCTCCCTTACCAGAGTTTCTATGAGA 838
 QY 121 ACGGCGCTCTCGGAATATTGATGCCCATATATAGGAGTTTATCTAAGGGAATCTCC 180
 DB 839 ATGACGGCTCCCGGAATATTGATGCCCATATGATGAGGAGTCTTTGTAAGGGAATCTCC 898
 QY 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAACTCTGCCACTCTTTGCAATG 240
 DB 899, ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATCACTCTGCCACTCTTTGCAATG 958
 QY 241 CATGCAATATCTATTATTGACAGGGAATATGATTAATCTAGTTGCTCTGGAGGACTT 300
 DB 959 CATGCAATATCTATTATTGACAGGGAATATGATTAATCTAGTTGCTCTGGAGGACTT 1018
 QY 301 GGAGCCACTGCTGCTGCACTTACTTCAACCATACAGTATGCTGATGGGGTGAAT 360
 DB 1019 GGAGTCACTGCTGCTGCACTTACTTCAACCATACAGTATGCTGATGGGGTGAAT 1078
 QY 361 CAAGGTGAGGCAAGAGAAAAAAGTAAGTAAGCAATCTCCCAACTGACCCGGGGACAT 420
 DB 1079 CAAGTACAGCAAGAGAAAAAAGTAAGTAAGCAATCTCCCAACTGACCCGGGGTACAT 1138
 QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAATATACATGAAACCCCTCCGTACC 480
 DB 1139 GGCACCTCTAGCCCTCAAAAGGACTAGTCTCTCAAAATATACATGAAACCCCTCCGTACC 1198
 QY 481 CATACTGCGCTGGTGGAGCTATTTAATACCACTCTGCTCGGCTCCATGAGGTCTCAGCC 540
 DB 1199 CATACTGCGCTGGTGGAGCTATTTAATACCACTCTGCTCGGCTCCATGAGGTCTCAGCC 1258
 QY 541 CAAACCCCTACTAATGTTGGATGCTGCTCCCTCGCACTTCAGGCCATATCAATTTCAATC 600

DB 1259 CAAAACCCCTACTAATGTTGGATATGCTCCCTCCCTGAACTTCAGGCCATATGTTTCAATC 1318
 QY 601 CCTGTTCTGTGAACAATGGAACTTCAGACAGAAATAAACAACAATTCCTGTTTGTAGTA 660
 DB 1319 CCTGTACCTGAAACAATGGAACTTCAGACAGAAATAAACAACAATTCCTGTTTGTAGTA 1378
 QY 661 GGACCTCTGTTTCCAACTGGAATACCCATACCTCAAACTCAGCTGCTGCTGCTGCTGCTGCT 720
 DB 1379 GGACCTCTGTTTCCAACTGGAATACCCATACCTCAAACTCAGCTGCTGCTGCTGCTGCTGCT 1438
 QY 721 AGCAATATATAGACACACACACAGCTCCCAATGATCAGGTGGGTGAACACCTCCCAACAGA 780
 DB 1439 AGCAATATATAGACACACACACAGCTCCCAATGATCAGGTGGGTGAACACCTCCCAACAGA 1498
 QY 781 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 1499 ATAGTCTGCTACCTCCCTCAGGAATATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1558
 QY 841 AATGGCTCTTCAGAATCTATGCTGCT 900
 DB 1559 AATGGCTCTTCAGAATCTATGCTGCT 1618
 QY 901 ACTGAACAGATTTATACAACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 1619 ACTGAACAGATTTATACAACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1678
 QY 961 CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGAGTGTAGGAGTGTAGGAGTGTAGGAGTGTAGGAGT 1020
 DB 1679 CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGAGTGTAGGAGTGTAGGAGTGTAGGAGTGTAGGAGT 1738
 QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATTAATGGTGACATGGAACAG 1080
 DB 1739 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATTAATGGTGACATGGAACAG 1798
 QY 1081 GTCACCTGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 DB 1799 GTCGCGGACTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1858
 QY 1141 CAAATCGAAGAGCTTTAGACTTGTAACTGCTAACCGCAAGAGGGGGAACCTGTTTATTTTAA 1200
 DB 1859 CAAATCGAAGAGCTTTAGACTTGTAACTGCTAACCGCTGAAGAGGGGGAACCTGTTTATTTTAA 1918
 QY 1201 GGAGAAAGAGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTAA 1230
 DB 1919 GGAGAAAGAGCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTAA 1948

Search completed: February 21, 2005, 07:45:51
 Job time : 776.929 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:54:34 ; Search time 260.948 seconds
(without alignments)
9286.612 Million cell updates/sec

Title: US-09-319-156B-9
Perfect score: 1481
Sequence: 1 agggccctccctatcatcac.....gtcttacaatggaacccca 1481

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	1481	4	US-08-979-847B-105
2	1335.4	90.2	2946	3	Sequence 105, Appli
3	1333.8	90.1	2763	4	Sequence 3, Appli
4	1333.8	90.1	8523	4	Sequence 5640, Ap
5	1333.8	90.1	13537	4	Sequence 21, Appl
6	1277.4	86.3	145320	4	Sequence 17382, A
7	302	20.4	1329	4	Sequence 15858, A
8	173.8	11.7	542	1	Sequence 108, Appl
9	173.8	11.7	542	1	Sequence 48, Appl
10	161.6	10.9	601	4	Sequence 2, Appli
11	160.8	10.9	190078	4	Sequence 184496,
12	160.8	10.9	190078	4	Sequence 12707, A
13	160.8	10.2	601	4	Sequence 17026, A
14	150.8	10.2	601	4	Sequence 202250,
15	150.8	10.2	77772	4	Sequence 202251,
16	150.8	10.2	77997	4	Sequence 17417, A
17	145.6	9.8	1722	3	Sequence 12249, A
18	145.6	9.8	1722	3	Sequence 58, Appl
19	145.6	9.8	1722	4	Sequence 54, Appl
20	144.6	9.8	2908	4	Sequence 249, App
21	142.4	9.6	8399	4	Sequence 26, Appl
22	131.2	8.9	1859	3	Sequence 46, Appl
23	131.2	8.9	1859	4	Sequence 46, Appl
24	131.2	8.9	1859	4	Sequence 42, Appl
25	130.2	8.8	1001	4	Sequence 406, App
26	116.8	7.9	1704	1	Sequence 1, Appli
27	116.8	7.9	21526	4	Sequence 14685, A

c	28	116.6	7.9	601	4	US-09-949-016-184497	Sequence 184497,
	29	115.2	7.8	7713	4	US-09-573-080A-30	Sequence 30, Appl
	30	115	7.8	7392	4	US-09-573-080A-44	Sequence 44, Appl
c	31	114.8	7.8	601	4	US-09-949-016-184495	Sequence 184495,
	32	114.8	7.8	80246	3	US-09-078-294-4	Sequence 4, Appli
	33	114.2	7.7	2518	3	US-09-011-745-1	Sequence 1, Appli
	34	114.2	7.7	5865	3	US-09-011-745-8	Sequence 8, Appli
	35	112.6	7.6	6529	4	US-09-573-080A-28	Sequence 28, Appl
	36	112.2	7.6	4776	4	US-09-555-352-1	Sequence 1, Appli
	37	106.4	7.2	2809	4	US-09-620-312D-171	Sequence 171, App
	38	104.8	7.1	2342	4	US-09-904-615-12	Sequence 12, Appl
c	39	103.4	7.0	80355	4	US-09-949-016-12735	Sequence 12735, A
c	40	103.4	7.0	80357	4	US-09-949-016-13572	Sequence 13572, A
c	41	96.6	6.5	46698	4	US-09-949-016-17323	Sequence 17323, A
c	42	96.6	6.5	54382	4	US-09-949-016-12139	Sequence 12139, A
	43	90.8	6.1	80595	3	US-09-078-294-3	Sequence 3, Appli
	44	89	6.0	464	3	US-09-120-653D-6	Sequence 6, Appli
	45	89	6.0	3910	3	US-09-120-653D-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-979-847B-105
; Sequence 105, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUXE, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE, PLC

STREET: P.O. BOX 19928

CITY: ALEXANDRIA

STATE: VA

COUNTRY: USA

ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-No. 6582703-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BERRIDGE, WILLIAM P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 39046A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 105:

SEQUENCE CHARACTERISTICS:

LENGTH: 1481 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 105:

US-08-979-847B-105

Query Match 100.0%; Score 1481; DB 4; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
DB |||||
QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
DB |||||
QY 61 GCACCCCTCCCTGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB |||||
QY 61 GCACCCCTCCCTGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB |||||
QY 121 AGCGGCTTCTGGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB |||||
QY 121 AGCGGCTTCTGGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB |||||
QY 181 ACCTTCACGCCCACACCCATATGCCGCACTGCTATACTCTGCCACTCTTTGCAATG 240
DB |||||
QY 181 ACCTTCACGCCCACACCCATATGCCGCACTGCTATACTCTGCCACTCTTTGCAATG 240
DB |||||
QY 241 CATGCAAACTACTTATTTGGACAGGGAATAATGATTAATCTCTAGTTGTCTGGAGACTT 300
DB |||||
QY 301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACAGTATGCTCTGATGGGGTGAATT 360
DB |||||
QY 301 GGAGCCACTGTCTGTTGGACTTACTTCACCCATACAGTATGCTCTGATGGGGTGAATT 360
DB |||||
QY 361 CRAAGTCAGCAAGCAAGAAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
DB |||||
QY 361 CRAAGTCAGCAAGCAAGAAAACAAAGTAAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
DB |||||
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
DB |||||
QY 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACC 480
DB |||||
QY 481 CATACTCGCTGTGAGCCCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB |||||
QY 481 CATACTCGCTGTGAGCCCTTATTAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB |||||
QY 541 CAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCCTCTCAGGCCATACATTTCAATC 600
DB |||||
QY 541 CAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCCTCTCAGGCCATACATTTCAATC 600
DB |||||
QY 601 CTGTTTCTGGAACAATGGAACAACCTTCAGCACAGAAATAACACCACTTCGTTTGTATGTA 660
DB |||||
QY 601 CTGTTTCTGGAACAATGGAACAACCTTCAGCACAGAAATAACACCACTTCGTTTGTATGTA 660
DB |||||
QY 661 GGACCTCTGTTTCCCAATCTGGAATAACCCATACCTCAAACTCACTGTTGTAATTTT 720
DB |||||
QY 661 GGACCTCTGTTTCCCAATCTGGAATAACCCATACCTCAAACTCACTGTTGTAATTTT 720
DB |||||
QY 721 AGCAATACTATAGACACAAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
DB |||||
QY 721 AGCAATACTATAGACACAAACAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
DB |||||
QY 781 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTTGTTACCTCAGCCTATCATTTGTTG 840
DB |||||
QY 781 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTTGTTACCTCAGCCTATCATTTGTTG 840
DB |||||
QY 841 AATGGCTCTTCAGAACTATGCTTCTCTCAATCTTCTAGTGCCCTATGACCACTCTAC 900
DB |||||
QY 841 AATGGCTCTTCAGAACTATGCTTCTCTCAATCTTCTAGTGCCCTATGACCACTCTAC 900
DB |||||
QY 901 ACTGAACAAGATTTATACATCATGTGTAACCTTAAGCCCAACAAAGAGTACCATT 960
DB |||||
QY 901 ACTGAACAAGATTTATACATCATGTGTAACCTTAAGCCCAACAAAGAGTACCATT 960
DB |||||
QY 961 CTTCCCTTTGTTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020
DB |||||
QY 961 CTTCCCTTTGTTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGGCATTGGCAGTATC 1020
DB |||||

RESULT 2

US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R. A.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Mi, Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B-AJ172A
; CURRENT APPLICATION NUMBER: US/09/175.928A
; CURRENT FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2946
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-175-928-3

Query Match 90.2%; Score 1335.4; DB 3; Length 2946;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTTCTTACCCCTTTTCGCTCTCACT 60
DB |||||
QY 928 ATGGCCCTCCCTTATCATTTTCTTCTTACTGTTCTTCTTACCTTTTCGCTCTCACT 987
DB |||||
QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB |||||
QY 988 GCACCCCTCCATGCGCTGTATGACAGTAGTCTCCCTTACCAAGAGTTTCTATGGAGA 1047
DB |||||
QY 121 AGCGGCTTCTGGAATAATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB |||||

Db 1048 ATGCAGCGTCCGGAAATATTGATGCCCCATCGTATAGGAGTCTTTTCTAAGGGAACCCCC 1107
Qy 181 ACCTTCACCTGCCACACCATATATGCCCCGCAACTGCTATAAATCTTCCCACTCTTTGCGATG 240
Db 1108 ACCTTCACCTGCCACACCATATATGCCCCGCAACTGCTATAAATCTTCCCACTCTTTGCGATG 1167
Qy 241 CATGCAAAATACTCATTATTGGACAGGGAATAATGATTAAATCCCTAGTGTCTCTGGAGGACTT 300
Db 1168 CATGCAAAATACTCATTATTGGACAGGGAATAATGATTAAATCCCTAGTGTCTCTGGAGGACTT 1227
Qy 301 GGAGCCACTGTCTGTGTGGACTTACTTCAACCCATACCAAGTATGCTGATGGGGGTGAAAT 360
Db 1228 GGAGTCACTGTCTGTGTGGACTTACTTCAACCCATGCTGATGCTGATGGGGGTGAGATT 1287
Qy 361 CAAGGTGAGCAAGAGAAAACAAGTAAAGGAGCAATCTCCAACTGACCCGGGACAT 420
Db 1288 CRAAGTCAGCAAGAGAAAACAATGTAAGAAGTAAATCTCCCACTCAACCCGGGTACAT 1347
Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTCTCTCAAAACTACATGAAACCCCTCCGTACC 480
Db 1348 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 1407
Qy 481 CATACTCGCTGTGTGAGCCCTATTAAATACCAACCTCACTGGGTCCATGAGGTCTCAGCC 540
Db 1408 CATACTCGCTGTGTGAGCCCTATTAAATACCAACCTCACTGGGTCCATGAGGTCTCAGCC 1467
Qy 541 CAAAACCCCTACTAATCTGTGTGGATGTGGTCCCTCCCTGCACCTTCAGGCAATACATTTCAATC 600
Db 1468 CAAAACCCCTACTAATCTGTGTGGATGTGGTCCCTCCCTGCACCTTCAGGCAATATGTTTCAATC 1527
Qy 601 CCTGTTCTCTCAACAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
Db 1528 CCTGTACTGAAACAAATGGAACAACTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 1587
Qy 661 GGACCTCTTGTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCACTCTGTGTAAATTT 720
Db 1588 GGACCTCTTGTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCACTCTGTGTAAATTT 1647
Qy 721 AGCAATACTATAGACACACAGCTGCCAATGATCGAGTGGGTAAACACCTCCACACGA 780
Db 1648 AGCAATACTATAGACACACAGCTGCCAATGATCGAGTGGGTAAACACCTCCACACAA 1707
Qy 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGACCTCAGCCTATCATTTGTTG 840
Db 1708 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTGTGACCTCAGCCTATCGTTGTTG 1767
Qy 841 AATGGCTCTTCAGAACTATGTGCTCTCTCAATTTAGTGTGCCCCCTATGACCATCTAC 900
Db 1768 AATGGCTCTTCAGAACTATGTGCTCTCTCAATTTAGTGTGCCCCCTATGACCATCTAC 1827
Qy 901 ACTGAAACAAGATTATACAATCTATGCTGTACCTTAAGCCCCCAACAACAAGATACCCATT 960
Db 1828 ACTGAAACAAGATTATACAATCTATGCTGTACCTTAAGCCCCCAACAACAAGATACCCATT 1887
Qy 961 CTTCTCTTTGTTATCAGCAGGAGTGTAGGAGACTAGTGTACTGTGCTTGGCAGTATC 1020
Db 1888 CTTCTCTTTGTTATGAGGAGGAGTGTAGGAGACTAGTGTACTGTGCTTGGCAGTATC 1947
Qy 1021 ACAACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGTTGATGGAACAG 1080
Db 1948 ACAACCTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGTTGATGGAACAG 2007
Qy 1081 GTCACTGATCTCCTGGTCACTTGCAGAGATCACTTAACCTCCCTAGCAGCAGTATCTT 1140
Db 2008 GTCCGCCAGTCCCTGGTCACTTGCAGAGATCACTTAACCTCCCTAGCAGCAGTATCTT 2067
Qy 1141 CAAATCGAGAGCTTTTGTAGCTTCTAAGCCGCAAAAGAGGGGAACTGTTTATTTTAA 1200
Db 2068 CAAATCGAGAGCTTTTGTAGCTTCTAAGCCGCTGAAGAGGGGAACTGTTTATTTTAA 2127
Qy 1201 GGAGAAGAACGCTGTATTATGTTAAATCAATCCAGAAATTTGTCACTGAGAAAGTTTAAAGAA 1260
Db 2128 GGGGAAGAAATGCTGTTATTATGTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAA 2187

Qy 1261 ATTCCAGATCGAATACAAATGTAGACGAGGAGCTTCAAAAACACCGAAGCGTGGGCGCTC 1320
Db 2188 ATTCCAGATCGAATACAAATGTAGACGAGGAGCTTCAAAAACACCGAAGCGTGGGCGCTC 2247
Qy 1321 CTCAGCCAAATGATGATGCCCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 2248 CTCAGCCAAATGATGATGCCCTGGTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTG 2307
Qy 1381 TTACTCTCTTTGGACCCCTGTATCTTTAACTCCCTGTTTAAAGTGTCTCTTCCAGAAAT 1440
Db 2308 CTACTCTCTTTGGACCCCTGTATCTTTAACTCCCTGTTTAACTCTTTGTCTCTTCCAGAAAT 2367
Qy 1441 GAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAACCCCA 1481
Db 2368 GAAGCTGTAAAGCTACAAATGGAACCCCAAGATGCAAGTCCAA 2408

RESULT 3

US-09-949-016-5640
; Sequence 5640, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5640
; LENGTH: 2763
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5640

Query Match 90.1%; Score 1333.8; DB 4; Length 2763;
Best Local Similarity 93.8%; Pred. NO. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATCTTTTCTCTTTACTGTTCTTACCCCTTTTCCCTCTCACT 60
Db 763 ATGGCCCTCCCTTATCATCTTTTCTCTTTACTGTTCTTTTACCCCTCTTTCACCTCACT 822
Qy 61 GCACCCCTCCATGCTGTACAAACAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 120
Db 823 GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTTTACCAAGAGTTTCTATGAAGA 882
Qy 121 ACCGCGCTTCTCGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAATCC 180
Db 883 ATCAGCGTCCCGAAATATTGATGCCCATCGTATAGGAGTCTTTCTAAGGAAACCCCC 942
Qy 181 ACCTTCACTGCCCAACCCATATGCGCCGCACTGCTATAAATCTGCACTCTTTGCAATG 240
Db 943 ACCTTCACTGCCCAACCCATATGCGCCGCACTGCTATAAATCTGCACTCTTTTGCATG 1002
Qy 241 CATGCAAAATCACTATTATTGGACAGGAAATGATTAATCTAGTGTCTCTGGAGGACTT 300
Db 1003 CATGCAAAATCACTATTATTGGACAGGAAATGATTAATCTAGTGTCTCTGGAGGACTT 1062
Qy 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCCATACCAAGTATGTCTGATGGGGTGAATT 360
Db 1063 GGAGTCACTGTCTGTTGGACTTACTTCAACCCAACTGATGTCTGATGGGGTGGAGTT 1122
Qy 361 CAAGGTGAGCAAGAGAAAACAAGTAAAGAGCAATCTCCCACTGACCCCGGACAT 420

QY 721 AGCAATACCTATAGACACACACACAGCTCCCAATGCATCAGGTGGGTAAACACCTCCACAGA 780
DB 10256 AGCAATACCTATAGACACACACACACTCCCAATGCATCAGGTGGGTAACTCTCCACACAA 10315
QY 781 ATAGTCTGCTACCTACCTCAGGAATATTTTTGTCTGTGTGTAACCTCAGCCATCATTTGTTG 840
DB 10316 ATAGTCTGCTACCTACCTCAGGAATATTTTTGTCTGTGTGTAACCTCAGCCATCATTTGTTG 10375
QY 841 AATGGCTCTTCAGAACTATGTGCTCCCTCATTTAGTGGCCCTATGACCATCTAC 900
DB 10376 AATGGCTCTTCAGAACTATGTGCTCCCTCATTTAGTGGCCCTATGACCATCTAC 10435
QY 901 ACTGAACAAGATTTATACAAATCATGTCTGTAACCTTAAGCCCCCAACAAAGAGTACCCATT 960
DB 10436 ACTGAACAAGATTTATACAGTTATGTCTATCTAGCCCGCGCAACAAAGAGTACCCATT 10495
QY 961 CTTCTCTTTGTTATCAGACGAGGTGCTAGGACACTAGGTACTGGCAATGGCGATATC 1020
DB 10496 CTTCTCTTTGTTATAGGACGAGGTGCTAGGTGCACTAGGTACTGGCAATGGCGATATC 10555
QY 1021 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
DB 10556 ACACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAATGGGACATGGAACAG 10615
QY 1081 GTCACTGACTCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
DB 10616 GTCCGCGACTCCCTGGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 10675
QY 1141 CAAATTCGAAGAGCTTTAGACTGTCTAACCGCCAAAGAGGGGAACTGTGTTATTTTAA 1200
DB 10676 CAAATTCGAAGAGCTTTAGACTGTCTAACCGCTGAAAGAGGGGAACTGTGTTATTTTAA 10735
QY 1201 GGAGAAAGAGCTGTTATTTATGTTAATCAATCAGAAATGTCACAGAAAGTTAAAGAA 1260
DB 10736 GGAGAAAGAGCTGTTATTTATGTTAATCAATCCGGAATCGTCACAGAAAGTTAAAGAA 10795
QY 1261 ATTGAGATCGAATCAATGTTAGAGCAGAGAGCTTCAAAACACCGAGCTGGGCGCTC 1320
DB 10796 ATTGAGATCGAATCAATGTTAGAGCAGAGAGCTTCAAAACACCGAGCTGGGCGCTC 10855
QY 1321 CTCAGCAATGAGTCCCTGGGTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 10856 CTCAGCAATGAGTCCCTGGATCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG 10915
QY 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCCCTGTTAAAGTTGTTCTCTCCAGAAAT 1440
DB 10916 CTACTCTCTTTGGACCTGTATCTTTAACTCCCTGTTAAAGTTGTTCTCTCCAGAAAT 10975
QY 1441 GAAGCTGTAAGCTACAGATGGTCTTTACAAATGGAAACCCCA 1481
DB 10976 GAAGCTGTAAGCTACAAATGGAGCCCAAGATGCAGTCCAA 11016

RESULT 6
US-09-949-016-15858/c
; Sequence 15858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15858

; LENGTH: 145320
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15858
Query Match 86.3%; Score 1277.4; DB 4; Length 145320;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 81; Indels 33; Gaps 4;
QY 1 ATGGCCCTCCCTATCATACATTTTCTCTTACTGTCTTACCCCTTTCGCTCTCACT 60
DB 94027 ATGGCCCTCCCTATCATATTTTCTCTTACTGTCTTACCCCTTTCGCTCTCACT 93968

QY 61 GCACCCCTCCCTGCTGTACAAACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 93967 GCACCCCTCCCTGCTGTACAAACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 93908
QY 121 ACGGGCTTCTGGAATAATTTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
DB 93907 ATGGGGCTTCCAGAAATATTTGATGCCCATCAATAGGAGTTTACCTTAAAGGAAACTCC 93848
QY 181 ACCTTCAGTCCCAACCCATATGCCCGCAACTGTATACTCTGCCACTCTTTGCGATG 240
DB 93847 ACCTTCAGTCCCAACCCATATGCCCGCAACTGTATACTCTGCCACTCTTTGCGATG 93788
QY 241 CATGCAATATCTCATTTATTGGACAGGGAAATGATTAATCTCTAGTGTCTCTGGAGACTT 300
DB 93787 CATGCAATATCTCATTTATTGGACAGGGAAATGATTAATCTCTAGTGTCTCTGGAGACTT 93728
QY 301 GGAGCACTGTCTGTGGAATCTTTACCCATACAGTATGTCTGATGGGGTGAATTT 360
DB 93727 GGAGCACTGTCTGTGGAATCTTTACCCATACAGTATGTCTGATGGGGTGAATTT 93668
QY 361 CAAGGTGAGCAAGAGAAACAGTAAGGAAGCAATCTCCCACTGACCCGGGAGCAT 420
DB 93667 CAAGGTGAGCAAGAGAAACAGTAAGGAAGCAATCTCCCACTGACCCGGGAGCAT 93608
QY 421 AGCACCCTAGCCCCACAAAGGACTAGTTCTCTC-AAAACTACATGAACCCCTCGGTAC 479
DB 93607 AGCACCCTAGCCCCACAAAGGACTAGTTCTCTC-AAAACTACATGAACCCCTCGGTAC 93548
QY 480 CCAATCTCGCTGTGAGCCCTATTAAATACACCCCTCACTGGCTCCATGAGTCTCAGC 539
DB 93547 CCAATCTCGCTGTGAGCCCTATTAAATACACCCCTCACTGGCTCCATGAGTCTCAGC 93488
QY 540 CCAAAACCTACTAACTGTGATGTGCTCCCTGCACTTTCAGGCAATGATTTCAAT 599
DB 93487 CCAAAACCTACTAACTGTGATGTGCTCCCTGCACTTTCAGGCAATGATTTCAAT 93428
QY 600 CCCTGTTCTTGAAACAAATGGAACAACTTCAGCAACAGAAATAAACACCACTTCGTTTGTAGT 659
DB 93427 CCCTGTTCTTGAAACAAATGGAACAACTTCAGCAACAGAAATAAACACCACTTCGTTTGTAGT 93368
QY 660 AGGACCTCTGTTTCCAAATCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATTT 719
DB 93367 AGGACCTCTGTTTCCAAATCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATTT 93308
QY 720 TAGCAATATAGACAAACAGCTCCCAATGCAATGAGTGGGTAAACACCTCCACACAG 779
DB 93307 TAGCAATATAGACAAACAGCTCCCAATGCAATGAGTGGGTAAACACCTCCACACAG 93248
QY 780 AATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCATCATTTGTTT 839
DB 93247 AATAGTCTGCTTACCCCTCAGGAATATTTTGTCTGTGTGTAACCTCAGCCATCATTTGTTT 93188
QY 840 GAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTTAGTGGCCCTATGACCATCTA 899
DB 93187 GAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTTAGTGGCCCTC-ATGACCATTTA 93129
QY 900 CACTGAAACAGATTTTATCAATCATGTCTAAGCCCAACAAAGAGTACCCAT 959
DB 93128 CACTGAAACAGATTTTATCAATCATGTCTAAGCCCAACAAAGAGTACCTCAT 93069

QY 960 TCTTCCTTTGTTTATCAGACGAGGAGTCTAGCAGACTAGTAGTCTGCGATTGGCAGTAT 1019
Db 93068 TCTTCCTTTGTTTATCAGACGAGGAGTCTAGTCTGCGATTGGCAGTAT 93009
QY 1020 CACAACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATAATGGTGACATGAACA 1079
Db 93008 CACAACCTCTACTCAGTCTTACTACAACTATCTCAAGAACTCAATGGTGACATGAATG 92949
QY 1080 GGTCACTGACTCCTCGTCACTTCAAGATCAATCTTACTCCTAGCAGCAGTACTCT 1139
Db 92948 GGTGCGACTCCTCGTCACTTCAAGATCAATCTTACTCCTAGCAGCAGTACTCT 92889
QY 1140 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTATTTT 1199
Db 92888 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCTGAAAGAGGGGAACTGTTATTTT 92829
QY 1200 AGGAGAAGAACGCTG-----TTATATGTTTAATCA 1229
Db 92828 AGGGAAGAAATGTTTATTATGTTATTTAGCGGAAGAAATGTTTATTATGTTAATCA 92769
QY 1230 ATCCAGAAATGTCACGTAGAGAAAGTTAAAGAAATTCGAGATCGAATCAATGTAGACAGA 1289
Db 92768 ATCTGAAATGTCACAGAGAAAGTTGAAGAAATTCGAGATTGAATCAACGTAGAACAGA 92709
QY 1290 GGAGCTTC-AAAACACCGAAGCTGGGGCTCTCCTCAGCCAAATGGATGCCCTGGGTTCTCC 1348
Db 92708 GGAGCTTC-AAAACACCGAAGCTGGGGCTCTCCTCAGCCAAATGGATGCCCTGGATCTCC 92649
QY 1349 CTTCTTTAGGACCTCTAGCAGCTCTAATATGTTTACTCTCTTTGGACCTGATCTTTA 1408
Db 92648 CTTCTTTAGGATCTCTAGCAGCTCTAATATGTTTACTCTCTTTGGACCTGATCTTTA 92589
QY 1409 ACTCCTTTGATGTTGCTCTTCCAGAAATGAGCTGTAAGCTTACAGATGCTCTTAC 1468
Db 92588 ACTCCTTTGATGTTGCTCTTCCAGAAATGAGCTGTAAGCTTACAGATGCTCTTAC 92529
QY 1469 AAATGGAAACCCA 1481
Db 92528 AAATGGAAACCCA 92516

RESULT 7
US-08-979-847B-108
; Sequence 108, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-NO. 6582703-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-08-979-847B-108
Query Match 20.4%; Score 302; DB 4; Length 1329;
Best Local Similarity 92.7%; Pred. No. 2.4e-87;
Matches 317; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1140 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTATTTT 1199
Db 1 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAACTGTTATTTT 60
QY 1200 AGGAGAAGAACGCTGTTATGTTAATCAATCCAGAAATTTGTCAGTGAAGAAAGTAAAGA 1259
Db 61 AGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTACTGAGAAAGTAAAGA 120
QY 1260 AATTGAGATCGAATCAATGTTAGACAGAGGAGCTTCAAAACACGAGAGCTGGGCGCT 1319
Db 121 AATTGAGATCGAATCAATGTTAGACAGAGGAGCTTCAAAACACTGCACCTGGGCGCT 180
QY 1320 CCTCAGCAATGATGCCCTGGGTTCTCCCTTTCTAGGACCTCTAGCAGCTCTAATATT 1379
Db 181 CCTCAGCAATGATGCCCTGGGTTCTCCCTTTCTAGGACCTCTAGCAGCTCTAATATT 240
QY 1380 GTTACTCTCTTTGGACCTGTTATCTTTAACTCTCTTTGTTAGTTGTTCTCTCCAGAAT 1439
Db 241 TTTACTCTCTTTGGACCTGTTATCTTTCACTCTCTTTGTTAGTTGTTCTCTCCAGAAT 300
QY 1440 TGAAGCTGTAAGCTACAGATGCTTTCATAAATGGAACCCCA 1481
Db 301 TGAAGCTGTAAGCTACAGATGCTTTCATAAATGGAACCCCA 342

RESULT 8
US-08-686-878A-48
; Sequence 48, Application US/08686878A
; Patent No. 5708157
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Bridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/686,878A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 542 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-686-878A-48

Query Match 11.7%; Score 173.8; DB 1; Length 542;
Best Local Similarity 91.5%; Pred. No. 7.7e-46;
Matches 184; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTATCATACACTTTTCTTTTACTGTTCTTACCTGTTCTTACCCCTTTGCTCTCACT 60
DB 342 ATGGCCCTCCCTATCATATTTTCTTTTACTGTTCTTTTACCTGTTCTTACCTCTTCTCACT 401

QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 402 GCACCCCTCCATGCTGTGTATGACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGGAGA 461

QY 121 ACGCGCTTCTGGAATAATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC 180
DB 462 ATGCAGGCTCCCGAAATATTGATGCCCATCATATAGAGTTTCTTCTAAGGGAACCTCC 521

QY 181 ACCTTCACCTGCCACACCCAT 201
DB 522 ACCTTCACCTGCCACACCCAT 542

RESULT 9
US-08-721-489-2
Sequence 2, Application US/08721489
Patent No. 5786465
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: LaValle, Edward
APPLICANT: Racie, Lisa
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,489
FILING DATE:

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 542 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-721-489-2

Query Match 11.7%; Score 173.8; DB 1; Length 542;
Best Local Similarity 91.5%; Pred. No. 7.7e-46;
Matches 184; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTATCATACACTTTTCTTTTACTGTTCTTACCTGTTCTTACCCCTTTGCTCTCACT 60
DB 342 ATGGCCCTCCCTATCATATTTTCTTTTACTGTTCTTTTACCTGTTCTTACCTCTTCTCACT 401

QY 61 GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 402 GCACCCCTCCATGCTGTGTATGACCAAGTAGTCTCCCTTACCAAGAGTTTCTATGGAGA 461

QY 121 ACGCGCTTCTGGAATAATTGATGCCCATCATATAGAGTTTATCTAAGGGAACCTCC 180
DB 462 ATGCAGGCTCCCGAAATATTGATGCCCATCATATAGAGTTTCTTCTAAGGGAACCTCC 521

QY 181 ACCTTCACCTGCCACACCCAT 201
DB 522 ACCTTCACCTGCCACACCCAT 542

RESULT 10
US-09-949-016-184496/c
Sequence 184496, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 184496
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-184496

Query Match 10.9%; Score 161.6; DB 4; Length 601;
Best Local Similarity 60.2%; Pred. No. 7.7e-42;
Matches 320; Conservative 1; Mismatches 205; Indels 6; Gaps 3;

QY 421 AGCACCCCTAGCCCTACAAAGGACTGTTCTCTCAAAACTCATGAACCCCTCGTACC 480
DB 532 AGTACTCCAGTCCATACAAGAAATTAGACCTTTCCAGGCTACAGGAAGCCCTTAACCT 473

QY 481 CATACTCGCTGGTGGAGCTATTTTATACCACCCCTCCTCGGCTCCATGAGGTCTCAGCC 540
DB 472 CATCTCATCTCTGGAGGCTCTTTTAAACACCCCTTTTACAGGAATACAGAGCCCTCTCT 413

QY 541 CAAACCCCTACTAATCTGTTGGATGCTCCCTCCCTGCTCAGGCATATCTTCAATC 600
 Db 412 GGTAACTCAACCACTGCTGGATGCTCTCCCTTGA- TTTTCAATGATGTCCAGTC 354
 QY 601 CTTGTTCTGAAACAATGGAACAATCTCAGCAGAGAAATAAACACCACTTCGGTTTGTAGTA 660
 Db 353 CTTGTTCTGAAACAATGGAACAATCTCAGCAGAGAAATAAACACCACTTCGGTTTGTAGTA 660
 QY 661 GGAACCTCTGTTTCCAACTGGAATAAACCACTCAGCAGAGAAATAAACACCACTTCGGTTTGTAGTA 720
 Db 296 AGTCCCATAGTACCAATTTACAGGCACACAGGCTCAAAATCTCAGTGCATAAACTTC 237
 QY 721 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCCAACAGTA 780
 Db 236 AGCATGCTGTCAATAGAACACCTCTGATGTCAGTCTGATATATAGTAACCCAGGT 177
 QY 781 ATAGTCTGCTCCTACCTCAGGAATAATTTTTTGTCTGTGTAACCTCAGCCTATCATTTGTTG 840
 Db 176 TTCACCTAACTTCAGGCA--TTTTTTTTTCACTGTGATTTACACAGGTTATGATGCTA 119
 QY 841 AATGGCTCTTCAGAAATCTATGTTCTTCTCTCAATTTCTAGTGCCCTATGACCATCTAC 900
 Db 118 AACAGCACTCTGAAAGAACTATGCTTCTCACTCTCTTTCTAGCAGCTCCCATGTCATATAT 59
 QY 901 ACTGAACAAGATTTATACATCATGCTGATCAATGCTTACCTTAAGCCCAACAAAGAG 952
 Db 58 ACTGAACGAGATTACAAAGTCTCTTTATACCCCAATCTCGCCACACATGAG 7

RESULT 11

US-09-949-016-12707
 ; Sequence 12707, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12707
 ; LENGTH: 190078
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(190078)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-12707

Query Match 10.9%; Score 160.8; DB 4; Length 190078;
 Best Local Similarity 59.8%; Pred. No. 5.2e-40;
 Matches 324; Conservative 0; Mismatches 212; Indels 6; Gaps 3;
 QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAATACATGAAACCCCTCCGTACC 480
 Db 95515 AGTACTCCAGTCCATCAAGAAATAGACCTTTCCAGGCTACAGGAAGCCCTTAACCTCT 95574
 QY 481 CATACTCGCTGGTGGCTATTTAATACCACTCTACTCGGCTCCATGAGGTCTCAGCC 540
 Db 95575 CATTTCTATCTCTGGAGCCTCTTTAACAACCACTTTACAGGAATACAGGAAGCCTCTCCT 95634
 QY 541 CAAACCCCTACTAATCTGTTGGATGCTCCCTCCCTGCACTTCAGGCCATATCTTCAATC 600

Db 95635 GGTAACTCAACCAACTGCTGGATGCTCTCCCTTGA- TTTTCAATGATGTCCAGTC 95693
 QY 601 CTTGTTCTGAAACAATGGAACAATCTCAGCAGAGAAATAAACACCACTTCGGTTTGTAGTA 660
 Db 95694 CTTGTTCTGAAACAATGGAACAATCTCAGCAGAGAAATAAACACCACTTCGGTTTGTAGTA 660
 QY 661 GGAACCTCTGTTTCCAACTGGAATAAACCACTCAGCAGAGAAATAAACACCACTTCGGTTTGTAGTA 720
 Db 95751 AGTCCCATAGTACCAATTTACAGGCACACAGGCTCAAAATCTCAGTGCATAAACTTC 95810
 QY 721 AGCAATACTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAAACACCTCCCAACAGTA 780
 Db 95811 AGCATGCTGTCAATAGAACACCTCTGATGTCAGTCTGATATATAGTAACCCAGGT 95870
 QY 781 ATAGTCTGCTCCTACCTCAGGAATAATTTTTTGTCTGTGTAACCTCAGCCTATCATTTGTTG 840
 Db 95871 TTCACCTAACTTCAGGCA--TTTTTTTTTCACTGTGATTTACACAGGTTATGATGCTA 95928
 QY 841 AATGGCTCTTCAGAAATCTATGTTCTTCTCTCAATTTCTAGTGCCCTATGACCATCTAC 900
 Db 95929 AACAGCACTCTGAAAGAACTATGCTTCTCACTCTCTTTCTAGCAGCTCCCATGTCATATAT 95988
 QY 901 ACTGAACAAGATTTATACATCATGCTGATCAATGCTTACCTTAAGCCCAACAAAGAGTACCATT 960
 Db 95989 ACTGAACGAGATTACAAAGTCTCTTTATACCCCAATCTCGCCACACATGAGCCTTTATT 96048
 QY 961 CT 962
 Db 96049 GT 96050

RESULT 12

US-09-949-016-17026
 ; Sequence 17026, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17026
 ; LENGTH: 190078
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(190078)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17026

Query Match 10.9%; Score 160.8; DB 4; Length 190078;
 Best Local Similarity 59.8%; Pred. No. 5.2e-40;
 Matches 324; Conservative 0; Mismatches 212; Indels 6; Gaps 3;
 QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAAATACATGAAACCCCTCCGTACC 480
 Db 95515 AGTACTCCAGTCCATCAAGAAATAGACCTTTCCAGGCTACAGGAAGCCCTTAACCTCT 95574
 QY 481 CATACTCGCTGGTGGCTATTTAATACCACTCTACTCGGCTCCATGAGGTCTCAGCC 540
 Db 95575 CATTTCTATCTCTGGAGCCTCTTTAACAACCACTTTACAGGAATACAGGAAGCCTCTCCT 95634
 QY 541 CAAACCCCTACTAATCTGTTGGATGCTCCCTCCCTGCACTTCAGGCCATATCTTCAATC 600

Db 95635 GGTAAATCAACCAACTGCTGGATGTGTCTCCCTTGCATTTTCAATGATATGTCCAGTC 95693
Qy 601 CCTGTTCTCGAACAATGGAACAATTCAGCAGAGAAATAAACACCACTTCGGTTTAGTA 660
Db 95694 CTTGTCCTTGGACGTGGAACTTATCCACCCGAGTACTAAGATCA---CCAATCAATC 95750
Qy 661 GGACCTCTTGTGTTTCCAAATCTGGAATAACCCATACCTCAAACTCACTGTGTGTAATTT 720
Db 95751 AGTCCCATAGTACCAAAATTTACCAGCCACACAGGCTCAAACTCTCACGTGCATAACTTC 95810
Qy 721 AGCATATATAGACACACAGGCTCCCAATCATCAGGTGGTAACACCTCCCAACGTA 780
Db 95811 AGCATGGCTGTCTAATAAGAACACCTTCTGATGTCACTGCTGTATATCAGTAACCCAGGT 95870
Qy 781 ATAGTCTGCTACCTCCAGCAATATTTTGTCTGTGGTACCTCAGCCCTATCATGTTTG 840
Db 95871 TTCACCTAATTCAGGCA--TTTGTCTGTGATTACACAGGTTATGATGCCTA 95928
Qy 841 AATGGCTCTTCAGAATCTATGTGCTTCTCTCACTTCTAGTCCCGCTATGACCATCTAC 900
Db 95929 AACAGCACTCTGAAGAACTATGCTTCTCACTCTCTTCTAGCACCTCCCATGTCCATAT 95988
Qy 901 ACTGAACAAGTTTATACATCATGTCGTAACCTTAAGCCCAACACAAAGAGGTACCCATT 960
Db 95989 ACTGAACGAGGTTACAAAGTCTCTTATACCCCAATCTCGGCACACATGAGCCTTATT 96048
Qy 961 CT 962
Db 96049 GT 96050

RESULT 13
US-09-949-016-202250
; Sequence 202250, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202250
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202250

Query Match 10.2%; Score 150.8; DB 4; Length 601;
Best Local Similarity 84.6%; Pred. No. 2.5e-38;
Matches 181; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
Qy 4 GCCCTCCCTTATCATATCTTTCTTTACTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 62
Db 297 GGCWCCCCCGCTAATATTTCTTTACTGTTGTTCTTACCCCTTCACTATCACTC 356
Qy 63 ACCCCCTCCATGCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGAAC 122
Db 357 ACCCCCTCCATGCGCTATACCTACAGTAGTCCCTTACCAAGAGTTTCTATGGAGAAT 416
Qy 123 GCGGCTTCTGGAATATTGATGCCCCATCATATAGAGTTTATCTAAGGGAACCTCCAC 182
Db 417 GCGGCTTCCCAAGAAATATTGATGCCCAATGTATAGGAGTTTTTCTAAAGGAACCCAC 476

Qy 183 CTTCACTGCCCAACCAATATGCGCGCAACTGC 216
Db 477 TTTCAACATCCACCAATATGCGCGCAACTGC 510
RESULT 14
US-09-949-016-202251
; Sequence 202251, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202251
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202251

Query Match 10.2%; Score 150.8; DB 4; Length 601;
Best Local Similarity 84.6%; Pred. No. 2.5e-38;
Matches 181; Conservative 0; Mismatches 32; Indels 1; Gaps 1;
Qy 4 GCCCTCCCTTATCATATCTTTCTTTACTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 62
Db 299 GGCWCCCCCGCTAATATTTCTTTACTGTTGTTCTTACCCCTTCACTATCACTC 358
Qy 63 ACCCCCTCCATGCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGAAC 122
Db 359 ACCCCCTCCATGCGCTATACCTACAGTAGTCCCTTACCAAGAGTTTCTATGGAGAAT 418
Qy 123 GCGGCTTCTGGAATATTGATGCCCCATCATATAGAGTTTATCTAAGGGAACCTCCAC 182
Db 419 GCGGCTTCCCAAGAAATATTGATGCCCAATGTATAGGAGTTTTTCTAAAGGAACCCAC 478
Qy 183 CTTCACTGCCCAACCAATATGCGCGCAACTGC 216
Db 479 TTTCAACATCCACCAATATGCGCGCAACTGC 512

RESULT 15
US-09-949-016-17417/c
; Sequence 17417, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17417
; LENGTH: 7772
; TYPE: DNA

```

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(77772)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17417

Query Match      10.2%; Score 150.8; DB 4; Length 77772;
Best Local Similarity 84.6%; Pred. No. 5.3e-37;
Matches 181; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 4 GCCTCCCTTATCATCTTTTCTTTACTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 62
Db 16203 GGCCTCCCTTATCATCTTTTCTTTACTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 62

QY 63 ACCCTTCCATGCTGCTGACACAGTACTGCTCCCTTACCAAGTTCCTATGAAGAAC 122
Db 16143 ACCCTTCCATGCTGCTGACACAGTACTGCTCCCTTACCAAGTTCCTATGGAGAA 16084

QY 123 GCGCTTCTTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCCAC 182
Db 16083 GCGCTTCTTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCCAC 16024

QY 183 CTTCACATGCCACACCCATATGCCCGCAACTGC 216
Db 16023 TTTCACATGCCACACCCATATGCCCGTGCATTC 15990

```

```

RESULT 16
US-09-949-016-12249/c
; Sequence 12249 Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

```

```

Query Match      10.2%; Score 150.8; DB 4; Length 77997;
Best Local Similarity 84.6%; Pred. No. 5.3e-37;
Matches 181; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 4 GCCTCCCTTATCATCTTTTCTTTACTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 62
Db 16207 GGCCTCCCTTATCATCTTTTCTTTACTGTTCTCTTA-CCCCCTTTCGCTCTCACTGC 16148

QY 63 ACCCTTCCATGCTGCTGACACAGTACTGCTCCCTTACCAAGTTCCTATGAAGAAC 122
Db 16147 ACCCTTCCATGCTGCTGACACAGTACTGCTCCCTTACCAAGTTCCTATGGAGAA 16088

QY 123 GCGCTTCTTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCCAC 182
Db 16087 GCGCTTCTTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCCAC 16028

```

```

QY 183 CTTCACATGCCACACCCATATGCCCGCAACTGC 216
Db 16027 TTTCACATGCCACACCCATATGCCCGTGCATTC 15994

RESULT 17
US-08-691-563C-58
; Sequence 58, Application US/08691563C
; Patent No. 6001987
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Frederic BESEME
; APPLICANT: Frederic BEDIN
; APPLICANT: Glauca PARANHOS-BACCALA
; APPLICANT: Florence KOMURIAN-PRADEL
; APPLICANT: Colette JOLIVET
; APPLICANT: Bernard MANDRAND
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLAC
; TIC PURPOSES
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,563C
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1722 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-691-563C-58

```

```

Query Match      9.8%; Score 145.6; DB 3; Length 1722;
Best Local Similarity 55.7%; Pred. No. 2.4e-36;
Matches 302; Conservative 0; Mismatches 234; Indels 6; Gaps 1;

QY 940 CACAACAAAGAGTACCCATTCTCTTTTGTGTTATCAGACGAGGTGCTAGCAGACTA 999
Db 346 CAGTGCAGAGCCATACAACTAATATCCCTATTATAGGTTAGGAATGGCTACTGCTACA 405

QY 1000 GGTACTGGCATTGGCAGTATCACAACCTCTACTCTCAGTTCTTACTACAACTATCTCAAGAA 1059
Db 406 GGAACCTGGAATAGCCGGTTTATCTTCTTACTTATCTACTACCATACACTCTCAAGAAT 465

QY 1060 ATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGCAGAGTCAACTTAAC 1119
Db 466 TTCTCAGACAGTTTGAAGAAATAATGAAATCTTCTTACTTTAATAATCCCAATTAGAC 525

QY 1120 TCCTTAGCAGCAGTAGTCTTCAAAATCGNAGAGCTTTAGACTTGTAAACCGCCAAAGA 1179
Db 526 TCTTTGGCAGCAATGACTCTTCCAAAACCGCCGAGCCACACCTCTCACTGTGAGAA 585

QY 1180 GGGGAACCTGTTTATTTTATAGGAGAAGACGCTGTTATTATGTTAATCAATCCAGAATT 1239

```

Db 586 GAGAGACTCGACCTCTTCTAGGGGAGAGTGTGTTTACCTAACCACTCAGGGATA 645
Qy 1240 GTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGATAGAGCAGAGGAGCTTCAA 1299
Db 646 GTACGAGATGCCACCTGGCATTTACAGGAAAGGGCTTCTGATATCAGACAATGCGCTTTCA 705
Qy 1300 AACACCGAAGCTGGGGCCCTCTCAGCCAATCGATGCCCTGGGTTCTCCCTTCTTAGGA 1359
Db 706 AACTCTTATACCAA-----CCTCTGGAGTTGGGCAACATGGCTTCTTCCATTTCTAGGT 759
Qy 1360 CCTCTAGCAGCTTAATATGTTTACTCTCTTTGGACCCCTGATCTTTAAACCTCCTTGT 1419
Db 760 CCATGGCAGCATCTTGCTGTACTACCTTTGGGCCCTGTATTTTAAAGCTTCTTGTC 819
Qy 1420 AAGTTGTCTCTTCAGAAATGGAAGCTGTAAGCTACAGATGGCTTACAAATGGAACCC 1479
Db 820 AAATTTGTTTCTCTAGGATCGAAGCCATCAAGCTACAGATGGTCTTACAAATGGAACCC 879
Qy 1480 CA 1481
Db 880 CA 881

RESULT 18
US-09-374-766-58
; Sequence 58, Application US/09374766
; Patent No. 6579526
; GENERAL INFORMATION:
; APPLICANT: Herve PERRON
; APPLICANT: Frederic BESEME
; APPLICANT: Frederic BEDIN
; APPLICANT: Glauca PARANHOS-BACCALA
; APPLICANT: Florence KOMURIAN-PRADEL
; APPLICANT: Collette JOLIVET
; APPLICANT: Bernard MANDRAND
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oliff & Berridge
; STREET: 700 South Washington Street, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/374,766
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/691,563
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Berridge, William P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 38588
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1722 base pairs
; TYPE: nucleotide
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-374-766-58

Query Match 9.8%; Score 145.6; DB 4; Length 1722;
Best Local Similarity 55.7%; Pred. No. 2.4e-36;
Matches 302; Conservative 0; Mismatches 234; Indels 6; Gaps 1;
Qy 940 CACAACAAAAGAGTACCCATTCCTCTTTGTTATAGAGCAGGAGTCTAGGCGAGCTA 999
Db 346 CAGTGACAGACATACAACTAATATCCCTATTTATAGGGTTAGGAATGGCTACTGCTACA 405
Qy 1000 GGTACTGGCATTCGACAGTATCACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAA 1059
Db 406 GGAATCTGGAATAGCGGTTTATCTACTTCAATATCTCTACTACCATACACTCTCAAGAAAT 465
Qy 1060 ATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTTCGAAAGTCAACTTAAC 1119
Db 466 TTCTCAGACAGTTTGCAGAAATATGAAATCTATCTTACTTTTACAATCCCAATTAGAC 525
Qy 1120 TCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGA 1179
Db 526 TCTTTGGCAGCATGACTCTCCAAAACCGCCGAGGCCACACCTCTCTCACTGCTGAGAAA 585
Qy 1180 GGGGGAACCTGTTTATTTTATAGGAGAAAGACGCTGTTATTTATGTTTAAATCCAGAATT 1239
Db 586 GGAGGACTCTGCACCTTCTTAGGGGAGAGTGTGTTTTTACCTAACCACTCAGGGATA 645
Qy 1240 GTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGACGAGAGGCTTCAA 1299
Db 646 GTACGAGATGCCACCTGGCATTTACAGAAAGGGCTTCTGATATCAGACAATGCCTTTCA 705
Qy 1300 AACACCGAAGCTGGGGCCCTCTCAGCCAATGGAATGCCCTGGGTTCTCCCTTCTTAGGA 1359
Db 706 AACTCTTATACCAA-----CCTCTGGAGTTGGGCAACATGGCTTCTTCCATTTCTAGGT 759
Qy 1360 CCTCTAGCAGCTTAATATGTTTACTCTCTTTGGACCCCTGATCTTTAAACCTCCTTGT 1419
Db 760 CCCATGGCAGCATCTTGTCTGTACTACCTTTGGGCCCTGTATTTTAAAGCTTCTTGTC 819
Qy 1420 AAGTTGTCTCTTCAGAAATGGAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCC 1479
Db 820 AAATTTGTTTCTCTAGGATCGAAGCCATCAAGCTACAGATGGTCTTACAAATGGAACCC 879
Qy 1480 CA 1481
Db 880 CA 881

RESULT 19
US-08-979-847B-54
; Sequence 54, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUBE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM: Floppy disk

CITY: Falls Church
STATE: VA
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/007,282B
FILING DATE: 19930121
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR., GERALD M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1644-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-007-282B-1

Query Match	7.9%	Score 116.8	DB.1	Length 1704
Best local Similarity	51.5%	Pred. No. 5.8e-27		
Matches 268	Conservative 0	Mismatches 252	Indels 0	Gaps 0
Qy	932	CTAAGCCCCACAACAAAGAGTACCCATTCTTCCTTTGTTATCAGACGAGGAGTGTAG	991	
Db	1178	CTAGGCGTCATNAGAGGGCAGTCCAGTTTATCCCCCTGCTTGTGGGTCTAGGATTTTCAG	1237	
Qy	992	GCAGACTAGGTACTGGCAATTGGCAGTATCAACCTCTACTCTAGTTCTCTACAAACTAT	1051	
Db	1238	GGGCTACACTTGTCTGGTGGACCGGGCTTGGGGTCTCCGTTTCACTTATCACAAGCTCT	1297	
Qy	1052	CTCAGAAATAAATGGTGCATGGACAGGTCACTGACTCCCTGGTCACCTTGCAGATC	1111	
Db	1298	CTAATCAATTGATTGAAGATGTCCAGGCTCTTTCAGGGACATCAATGACCTACAGGACC	1357	
Qy	1112	AACTTACTCCCTAGCAGCAGTAGTCTCTCAAATCGAAGAGCTTTAGACTTGTAAACCG	1171	
Db	1358	AGATTGACTCTCCGCTGAGTTGTCTTACAAATAGAGAGGGTTAGACCTATTGACTG	1417	
Qy	1172	CCAAAAGAGGGGAACCTGTTTATTTTTAGGAGAGAACGCTGTTATATGTTTAATCAAT	1231	
Db	1418	CCGAAACAAGGAGGAATATGTCTCGCACTCCAGGAGAAAGTGTGTTTTCAGCTAAACAAGT	1477	
Qy	1232	CCGAAATTGTCTGAGAGAAAGTTAAAGAAATTCGAGATCGAATACATATGATAGCAGAGG	1291	
Db	1478	CGGGTATTCGATCGTGACAAGATCCGAAAACCTCCAAGAGGACCTTATCCGAGAGAAACGTG	1537	
Qy	1292	AGCTTCAAAACACCGAAGCGCTGGGGCTCTCTCAGCCAATGGATGCCCTGGGTCTCCCCCT	1351	
Db	1538	CACTTGTAGCAACCCCTGTGGAGCGGCTTGAACGGCTTCCTTCCATATTTGTCACCT	1597	
Qy	1352	TCTTTAGGACCTCTAGCAGCTCTAATATTTGTTACTCTCTTTTGGACCTGTATCTTTAAC	1411	
Db	1598	TGTTTAGCCCCCTCTTTTGGGCTCATATTTGTTCTGACCTCGGCCCTGCATTTATGAAGA	1657	
Qy	1412	TCCTTTGTTAAGTTGTCTCTTCCAGAATTAAGCTGTAAA	1451	
Db	1658	CCCTGACTCGCATATATACATGACAAAAATTCAGGCAGTAAA	1697	

RESULT 27
US-09-949-016-14685/C
; Sequence 14685, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 14685
; LENGTH: 21526
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14685

```

RESULT 28
US-09-949-016-184497/c
; Sequence 184497, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

```

Fri Feb 25 16:26:32 2005

;; PUBLICATION INFORMATION:
;; AUTHORS: Jurka, J; Malchewicz, J; Milosavljevic, A
;; TITLE: Prototypic sequences for human repetitive DNA
;; JOURNAL: Journal of Molecular Evolution
;; VOLUME: 35
;; ISSUE: 4
;; PAGES: 286-291
;; DATE: 1992-10-
;; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
;; DATABASE ENTRY DATE: 1996-01-26
;; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-30

Query Match 7.9%; Score 116.6; DB 4; Length 601;
Best Local Similarity 58.5%; Pred. No. 3.5e-27;
Matches 257; Conservative 1; Mismatches 175; Indels 6; Gaps 3;

QY 524 TCCATGAGTCTCAGCCCAAAACCCCTACTAACTGTGGATGTGCTCCCTCCCTGCACTTCA 583
DB 599 TACAAGAGCCCTCTCTGCTATCCAACTGCTGGATGTCTCCCTTGCA-TTTT 541
QY 584 GGCATACATTTCAATCCCTGTCTGAACTGGAACAACTTCAGCACAGAAATAACA 643
DB 540 CAATGATGTCCAGTCCCTGTCCCTGGACAGTGGAACTTATCCACCCAGTACTAAACA 481
QY 644 CCACCTCCCTTTTAGTAGACCTCTGTTCCTCAATCTGGAAATAACCCATACCTCAACC 703
DB 480 TCA---CCAAATCAATCAGTCCCATAGTCCCAATTTACCGCCACACAGGCTCAATC 424
QY 704 TCACCTGTGTAAATTTAGCAATCTATAGACACCAACAGCTCCCAATCATCAGGTGG 763
DB 423 TCACGTGCATTAACCTCAGCATGGCTGTCAATAGAACACCTCTGATGTGCTGTA 364
QY 764 TAACACCTCCACACAAATAGTCTGCCTACCTCAGGAATAATTTTGTCTGTGTACT 823
DB 363 TATCAGTAACCCAGGTTTCACTAACTTCAGGCA--TTTTTTTTTTCATCTGTGATTACA 306
QY 824 CAGCCTATCATTTGTTGAATGCTCTTCAGAACTCTGCTCTCTCTCATCTTAGTGC 883
DB 305 CAGGTAATGATGCTTAAACAGACACTCTGAAAGAACTATGCTTCTCACTCTTCTAGCAC 246
QY 884 CCCTATGACCATCTACACTGAACAGATTTATACAAATCATGTCTGATACCTAAAGCCCA 943
DB 245 CTCCATGTCATATATATCTGAAGAGAGTACAAAGTCTCTTATACCCCAATCTGCGC 186
QY 944 ACAAAGAGTACCAATCT 962
DB 185 ACACATGAGCCTTTATTGT 167

RESULT 29
US-09-573-080A-30
; Sequence 30, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 7713
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(7713)
; OTHER INFORMATION: herve
; PUBLICATION INFORMATION:

;; PUBLICATION INFORMATION:
;; AUTHORS: Jurka, J; Malchewicz, J; Milosavljevic, A
;; TITLE: Prototypic sequences for human repetitive DNA
;; JOURNAL: Journal of Molecular Evolution
;; VOLUME: 35
;; ISSUE: 4
;; PAGES: 286-291
;; DATE: 1992-10-
;; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
;; DATABASE ENTRY DATE: 1996-01-26
;; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-30

Query Match 7.8%; Score 115.2; DB 4; Length 7713;
Best Local Similarity 54.0%; Pred. No. 5e-26;
Matches 259; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

QY 997 CTAGTACTGGATTGGCAGTATCACAACTCTACTCAGTTCTACTCAAACTATCTCAA 1056
DB 7018 CTCGTAATCTGGAATAGCAGTCAATTTCAACCTCTGTCAAGACCTTCGTAAGCTGTCTAAT 7077
QY 1057 GAAATAAATGCTGACATGGAACAGTCACTGACCTCCCTGGTCACTTGCAGATCAACTT 1116
DB 7078 GACTTCTCTGTAGCATCACAGATGTGCCAAACTTTATCAGTCTCCAGCTTAAGTT 7137
QY 1117 AACTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTCTTAACCCCAAA 1176
DB 7138 GACTCTTTAGCTGCACTGTCTCCAAACCGCTGAGGCTTGACTTACTCAGTCTGAA 7197
QY 1177 AGAGGGGAACTGTATTATTTTAGGAGAAAGCCTGTATTATTTAGTTAACTCAATCCAGA 1236
DB 7198 AAGAGGAGCTGTGTATATTTCTTAAATGAAGACTGTGTCTTTTACCTAAACAACTGGC 7257
QY 1237 ATTGTCAGTGAAGTTTAAAGAAATTCGAGATCGAATACAAATGATAGAGC---AGAGGAG 1293
DB 7258 CTGGTGTATGACACATAAAACTCAAGGAAGAGACCAAACTTGCACCAAGCA 7317
QY 1294 CTTCAAAACACCGACGCTGGGGCTCTCAGCCAAATGGATGCCCTGGGTTCTCCCTTC 1353
DB 7318 AGTAATATGTAATCCCTTGGGCACTCTCTAATTCATGCTCTGGGTCTCCCAATT 7377
QY 1354 TTAGACCTCTAGCAGCTCTAATATTTAGTCTCTCTTTGGACCTGTATCTTTAACCTC 1413
DB 7378 CTTAGTCTCTTAATACCAATTTTCTCTCTTTTATTCGACCTGTATCTTCCATTTA 7437
QY 1414 CTTGTTAAGTTGTCTCTTCCAGAAATTCAGAGCTGAAAGCTACAGATGGTCTTACAAATG 1473
DB 7438 GTTCTCAATTCATCCAAACTGTATCCAGGCCATCACAATCATTTTATACGACAAATG 7497

RESULT 30
US-09-573-080A-44
; Sequence 44, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 7392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(7392)
; OTHER INFORMATION: herve
; PUBLICATION INFORMATION:

AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
TITLE: Prototypic sequences for human repetitive DNA
JOURNAL: Journal of Molecular Evolution
VOLUME: 35
ISSUE: 4
PAGES: 286-291
DATE: 1992-10-
DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
DATABASE ENTRY DATE: 1996-01-26
DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-44

Query Match 7.8%; Score 115; DB 4; Length 7392;

Best Local Similarity 56.9%; Pred. No. 5.6e-26;
Matches 211; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 1045 AACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCTACTGACCTCCCTGGTGCACCTTG 1104

Db 6904 AATTATCTAATCAGCTAATTTCTGATGTACAAATCTTATCTAGCACCATACAAAGATCTG 6963

QY 1105 CAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTG 1164

Db 6964 CAAGATCAAGTAGACTCATTTAGCCGAAGTGTCTCCAGAACAGAGGGGGCTAGATCTA 7023

QY 1165 CTAACGCCAAAAGAGGGGAACTGTTTATTTTATTTAGGAGAGAACGCTGTTATATGTT 1224

Db 7024 CTTACAGCAGAACAGGAGGAACTGTTTATTTAGGAGAGAACGCTGTTATATGTT 7083

QY 1225 AATCAATCCAGAAATTTGCTACTGAGAAAGTTAAAGAAATTCAGATCGAATACAAATGAGA 1284

Db 7084 AACAGTCAAGGATTTGAGAGACAAATAAATAAACCCTTACAGAGAACTAGAAAGACGT 7143

QY 1285 GCGAGAGCTTCAAAACACCGAAGCGTGGGGCTCTCAGCCATCGATGCGCTGGTT 1344

Db 7144 AGAAAGATCTAGCTTCCCAACCCACTTTGCACTGGGCTTCAAGGGCTCTCCCTTACCTC 7203

QY 1345 CTCCTCTTCTAGGACCTCTAGCAGCTCTAATATTTTACTCTCTTTGACCTCTGATC 1404

Db 7204 CTGCTCTTCTTGGCCCTCTACTTACCTCTCTGCTCTTACTCACCATTGGCGCTGCATT 7263

QY 1405 TTTAACCTCTCT 1415

Db 7264 TTTAACCTCT 7274

RESULT 31

US-09-949-016-184495/c

; Sequence 184495, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 184495

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-184495

Query Match 7.8%; Score 114.8; DB 4; Length 601;

Best Local Similarity 61.8%; Pred. No. 1.4e-26;

Matches 215; Conservative 1; Mismatches 128; Indels 4; Gaps 2;

QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTTCTTCAAACTACATGAAACCCCTCCGTACC 480

Db 349 AGTACTCCAGTCCATACAGAAATTTAGACCTTTCCAGGCTACAGGAARCCCTTAACTCT 290

QY 481 CATACTCGCCTGGTGAGCCTATTTAATACCACCCCTCAGTCGGCTCCATGAGGTCTCAGCC 540

Db 289 CATTTCTCATCTCTGGAGCCTCTTTAAACACCCCTTTACAGGAATACAGAAGCCTCTCCT 230

QY 541 CAAAACCCCTACTAACTGTTGGATGTCCTCCCTGCACTTCAGGCGCATACATTTCAATC 600

Db 229 GGTATATCCAAACCAACCTGCTGGATGCTCTCCCTTGCA-TTTTCAATGATGTCCAGTC 171

QY 601 CCTGTTTCTGAAACAACTTACGACAGAAATAAACAACCACTTCCGTTTGTAGTA 660

Db 170 CCTGCTCCCTGGACAGTGAACCTTATCCACCCAGTACTAAACATCA---CCAATCAATC 114

QY 661 GGACCTCTGTTTCCCTGCTGGAATACCCATCTCAACCTCAACCTGCTGTAATAATTT 720

Db 113 AGTCCCATAGTCACCAATTTTACCAGCCACACAGGCCCTCAAACTCTCAGTGCAATAAACTTC 54

QY 721 AGCAATCTATAGACAAACCCAGCTCCCAATGATCAGGTGGGTAACA 768

Db 53 AGCATGGCTGTCATGAACACCTTCTGATGTCACTCTGTATATCA 6

RESULT 32

US-09-078-294-4

; Sequence 4, Application US/09078294

; Patent No. 6265211

; GENERAL INFORMATION:

; APPLICANT: Choo, Kong-Hong Andy

; APPLICANT: Du Sart, Desiree

; APPLICANT: Cancilla, Michael R.

; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE

; FILE REFERENCE: Davies Col

; CURRENT APPLICATION NUMBER: US/09/078,294

; CURRENT FILING DATE: 1998-05-13

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 80246

; TYPE: DNA

; ORGANISM: Nucleotide sequence of NC-contig

US-09-078-294-4

Query Match 7.8%; Score 114.8; DB 3; Length 80246;

Best Local Similarity 55.7%; Pred. No. 2.9e-25;

Matches 263; Conservative 0; Mismatches 202; Indels 7; Gaps 2;

QY 1010 TTGGCAGTATCAACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGGTG 1069

Db 58509 TTAACAGCGTAAATGATGATTTATCTACTACTCTCAGACTCTCTCAAGAGATTTCTCAGACA 58568

QY 1070 ACATGGAACAGGTCACTGACTCCCTGGTCACTTGCAGAGTCACTTAACTCCCTAGCAG 1129

Db 58569 GTTTGCAAAAAAGAACGAAATCTGTCTTACTCTACAAATCCCAATAGACTCTTTTGGCAG 58628

QY 1130 CAGTAGTCTTCAAAATCGAAGCTTTAGACTTGTAAACCGCCAAAGAGGGGGAACCT 1189

Db 58629 CAGTAGTCTTCAAAACCCGCTGAGGCTAGACTCTCTTACTGCTGAGAAAGAAAGATCT 58688

QY 1190 GTTTATTTTGGAGAGAAACGCTGTTTATTTATGTTAATCAATCCAGAAATTTGCTCAGTA 1249

Db 58689 GCAC-TTCTTAGGGGTAGAGTGTGTTTTTATCTACTAACAGTCAAGGATTAATAGATA 58747

QY 1250 AAGTTAAAGAAATTCGAGATCGAATAACAATGTTAGAGCAGAGGAGCTTCAAAAACCGAAC 1309

Db 58748 CCACCCAGTGTTTTACAGGAAAGGCTTCTGAAATCAGACAATGCTTTCAAACTCTTTATA 58807

QY 1310 GCTGGGGCCCTCTCAGCCAAATGGATGCCCTGGGTCTCCCTTCTTTAGGACCTTAGCAG 1369

Db 58808 CCAA-----CCTCTGGAGTTGGCGCAGATGGGCTTCTCCCTTTCTTAGGCTCTGTGACAG 58861

QY 1370 CTCTAATATGTTACTCTCTTTGACCCCTGTATCTTTAACTCCTCTTCTTAAGTTTGTCT 1429
Db 58862 CCATCTTGTCTAATAGTCGCAITTTGGCCCTGTATTTTAACTCTTTGTCAAATTTGTTT 58921
QY 1430 CTTCCAGAAATTGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAACCCCA 1481
Db 58922 CCTCTAGGATCGAGGCCATCAAGCTACAGATGATCTTCAAAATGTAACCCCA 58973

RESULT 33

US-09-011-745-1
; Sequence 1, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: RD114
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: n is any nucleotide
US-09-011-745-1

Query Match 7.7%; Score 114.2; DB 3; Length 2518;
Best Local Similarity 55.0%; Pred. No. 5.2e-26;
Matches 224; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1045 AAATCTATCTCAAGAAATAAATGGTGACATGGAACAGGTCTACTCTCCCTGGTCACCTTG 1104
Db 1834 AAATATATCCCATCAGTTAATATCTGATGTCCAAAGTCTTATCCGGTACCATACAAGATTTA 1893
QY 1105 CAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTG 1164
Db 1894 CAAGACGAGTAGACTCGTTAGCTGAAGTAGTTCTCCAAATAGGAGGGGACTGGACCTA 1953
QY 1165 CTAACCGCCAAAAGAGGGGAACTGTATTTTATAGGAGAGAAACGCTGTTATATGTT 1224
Db 1954 CTAAGGSCAGACAAGGAGGAAATTTGTTAGCCTTACAGAAAAATGCTGTTTATGCT 2013
QY 1225 AATCAATCCAGAAATGTCATCGAAGATTAAAGAAATTCGAGATCGAATACAAATGAGA 1284
Db 2014 AACAAATGCAAGAAATTTGAGAAAACAAATTAAGAAACCCCTACAGAAAGAAATTACAAAACGC 2073
QY 1285 GCAGAGAGCTTCAAAACACCGAAGCTGGGGCTCTCTCAGCCCAATGGATGCCCTGGGTT 1344
Db 2074 AGGGAAGCTGGCAACCAACCTCTCTGACCGGGCTGCAGGGCTTCTTCGGTACCTC 2133
QY 1345 CTCCTCTCTTAGGACCTTAGCAGCTCTAATATGTTTACTCTCTTTGGACCCCTGTATC 1404
Db 2134 CTACCTCTCTGGGACCCCTACTCACCTCTCTACTACTACTAACCATTGGGCCATCGGTT 2193
QY 1405 TTTAACCTCTCTTGAATTTGTTCTCTCCAGAAATGAACTGTAAA 1451
Db 2194 TTCAGTGCCTCATGGCCCTTCATTAATGATAGACTTAATGTTGACA 2240

RESULT 34

US-09-011-745-8
; Sequence 8, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; CURRENT FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3611)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3612)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3614)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3613)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3799)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3800)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3801)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3802)
; OTHER INFORMATION: n is any nucleotide
US-09-011-745-8

Query Match 7.7%; Score 114.2; DB 3; Length 5865;
Best Local Similarity 55.0%; Pred. No. 8.8e-26;
Matches 224; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 1045 AAATCTCTCAAGAAATAAATGGTGACATGGAACAGGTCTACTCTCCCTGGTCACCTTG 1104
Db 2692 AAATATATCCCATCAGTTAATATCTGATGTCCAAAGTCTTATCCGGTACCATACAAGATTTA 2751
QY 1105 CAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTG 1164
Db 2752 CAAGACCAAGGTAGACTCGTTAGCTGAAGTAGTTCTCCAAATAGGAGGGGACTGGACCTA 2811
QY 1165 CTAACCGCCAAAAGAGGGGAACTGTATTTTATAGGAGAAACGCTGTTTATTATGTT 1224

Db 2812 CTAACGCGAGAACAGGAGGAAATTTGTTAGCCTTACAGAAAATGCTGTTTTATGCT 2871
 Qy 1225 AATCAATCCAGAAATGCTACAGAGAAATTAAGAAATTCAGATCGAATACAAATGTAGA 1284
 Db 2872 AACAAGTCAGGAATTTGAGAAACAAATTAAGAACCCCTACAGAGAAATTTACAAACGC 2931
 Qy 1285 GCAGAGAGCTTCAAAACACCGAACGCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGGT 1344
 Db 2932 AGGGAAGCTTGGACCAACCCCTCTCTGGACCGGCTGCAGGCTTCTTTCGTTACCTC 2991
 Qy 1345 CTCCTCTTTAGACCTCTAGCAGCTCTAATATTTGTTACTCTCTTTGACCCCTGTATC 1404
 Db 2992 CTACCTCTCTGGACCCCTACTCACCTCTCTACTCATACTAAACCAATTTGGCCATCGGT 3051
 Qy 1405 TTTAACTCTCTGTTAGTTTGTCTCTCCAGAAATGAAGCTGTAA 1451
 Db 3052 TTCAGTCGCTCATGCGCTTCATTAATGATAGACTTAATGTTGTACA 3098

RESULT 35
 US-09-573-080A-28
 ; Sequence 28, Application US/09573080A
 ; Patent No. 6828097
 ; GENERAL INFORMATION:
 ; APPLICANT: JOAN, KNOLL
 ; APPLICANT: ROGAN, PETER
 ; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
 ; FILE REFERENCE: 30307
 ; CURRENT APPLICATION NUMBER: US/09/573.080A
 ; CURRENT FILING DATE: 2000-05-16
 ; NUMBER OF SEQ ID NOS: 479
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 28
 ; LENGTH: 6529
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: repeat region
 ; LOCATION: (1)...(6529)
 ; OTHER INFORMATION: herfth2
 ; PUBLICATION INFORMATION:
 ; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
 ; TITLE: Prototypic sequences for human repetitive DNA
 ; JOURNAL: Journal of Molecular Evolution
 ; VOLUME: 35
 ; ISSUE: 4
 ; PAGES: 286-291
 ; DATE: 1992-10-
 ; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
 ; DATABASE ENTRY DATE: 1996-01-26
 ; DATABASE ENTRY DATE: 1996-01-26
 US-09-573-080A-28

Query Match 7.6%; Score 112.6; DB 4; Length 6529;
 Best Local Similarity 54.3%; Pred. No. 3.1e-25;
 Matches 274; Conservative 0; Mismatches 224; Indels 7; Gaps 2;
 Qy 929 TACCTAAGCCCAACAAAGAGTACCCTATCTTCTCTTTTGTATCAGAGCAGGATGC 988
 Db 5839 TCCCTATCTGACAGCAGSATACTATCAACTCATACCTCTGTAGTAACCCCTCAGAAATA 5898
 Qy 989 TAGGAGACTAGTACTGGCAATGGCAGTATCACAACCTCTACTAGTTCTACTACAAC 1048
 Db 5899 CTACAGGAGTTGGAACTGGGATTTGGGATTAACCACTCCCTTCTTATTAACCAATCCC 5958
 Qy 1049 TATCTCAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCTCTGTCTACCTTGAAG 1108
 Db 5959 TCTCAAAGACTTCAAGATAGCTTTGGAGAGATAGCAAAATCAATCAACTCTCCNAT 6018
 Qy 1109 ATCAACTTAACTCCCTAGCAGCAGTAGTCTTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAA 1168
 Db 6019 CACAAATAAATCT-TAGCAGCAGTAGTCTTCTTCAAAACCGCAGAGGCTTTAGACCTGTTCA 6077

Qy 1169 CCGCCAAAGAGGGGAAACCTGTTATTTTATAGGAGAAACGCTGTTTATTTATGTTAATC 1228
 Db 6078 CAGCGAAAAAGAGAACTCTGCTTTTCTAGATGAAACAGTATTCCTTTTATCTTAACC 6137
 Qy 1229 AATCCAGAAATTCCTACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAG 1288
 Db 6138 AATTCGCGCTGATACAGGATCTGTAAAGAAAGACTAAAGACCGAGCACAATAAATAAG 6197
 Qy 1289 AGGAGCTTCAAAACACCGAACGCTGGGCTC-----CTCAGCCAAATGGATGCCCTGGG 1342
 Db 6198 AAAACGCTCTCTGATGGCCAGCGTGGCTCTCTGCTCTTTTAGTACCTGCTTTCCATGGC 6257
 Qy 1343 TTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATTTGTTATCTCTCTTTTGGACCTGTGA 1402
 Db 6258 TACTGCGCTCTCTAGGCTGCGCTAAACCAATTTTCTCTTTCTAGCATTTGGCCCTGTGC 6317
 Qy 1403 TCTTTAACTCTCTTGTAAAGTTTGT 1427
 Db 6318 TCTTACATCTCTTACCAGTTTTT 6342

RESULT 36
 US-09-555-352-1
 ; Sequence 1, Application US/09555352
 ; Patent No. 6544779
 ; GENERAL INFORMATION:
 ; APPLICANT: Cichutek, Klaus
 ; APPLICANT: Merges-Millitzer, Heike
 ; TITLE OF INVENTION: PSEUDO-TYPE RETROVIRAL VECTORS WITH
 ; TITLE OF INVENTION: MODIFIABLE SURFACE CAPSID PROTEINS
 ; FILE REFERENCE: 11692-005001
 ; CURRENT APPLICATION NUMBER: US/09/555,352
 ; CURRENT FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: PCT/DE98/03542
 ; PRIOR FILING DATE: 1998-11-27
 ; PRIOR APPLICATION NUMBER: DE 197 52 855.4
 ; PRIOR FILING DATE: 1997-11-28
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 4776
 ; TYPE: DNA
 ; ORGANISM: Murine leukemia virus
 US-09-555-352-1

Query Match 7.6%; Score 112.2; DB 4; Length 4776;
 Best Local Similarity 51.5%; Pred. No. 3.5e-25;
 Matches 258; Conservative 0; Mismatches 243; Indels 0; Gaps 0;
 Qy 951 AGTACCCATCTTCTCTTTTCTTATCAGAGCAGGAGTCTAGGAGACTAGGTACTGGCAT 1010
 Db 1216 AGTCCAGTTTATCCCTCTGCTTGTGGTCTAGGGATTTACAGGGCTACACTTGTGTTGG 1275
 Qy 1011 TGGCAGTATCACAACCTTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTTGA 1070
 Db 1276 AACGGGCTTGGGGTCTCCGCTTCACTTATCAAGCTCTCTAATCAATGTTGAAGA 1335
 Qy 1071 CATGGAAACAGTCACTGACTCCCTGGTCACTCTGCAAGATCAACTTAATCCCTTAGCAGC 1130
 Db 1336 TGTCAGGCTCTTTACGGGACCATCAATGACCTACAGGACAGATTGACTCCCTGGCTGA 1395
 Qy 1131 AGTAGTCTTTCAAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCAAAAGAGGGGAACTG 1190
 Db 1396 GGTTGCTTTACAAATAGAGAGAGGGTTAGACCTTATGACTTGCAGAACAGAGGAGAAATG 1455
 Qy 1191 TTTATTTTATAGGAGAAAGAACGCTGTTTATTTATGTTTAATCAATCCAGAAATTTGCTAGAA 1250
 Db 1456 TCTCGCACTCCAGGAGAGTGTGTTTACCGCTAACAGTCCGGTATCGTACGTGACAA 1515
 Qy 1251 AGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCGAAGC 1310
 Db 1516 GATCCGAAAACTCCAGAGGACCTTATCGAGAGAAAAACGTCGACTGTACGACAAACCCCT 1575

QY 1311 CTGGGCGCTCTCAGCCAAATGATGCCCTGGGTTCTCCCTTCTTAGACCTTAGCAGC 1370
 Db 1576 GTGGAGCGCTTGAACGGCTTCTTCCATATTTGCTACCTTGTGTAGGCCCTCTGTTGG 1635
 QY 1371 TCTATATTTGTTACTCCTCTTTGGACCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTC 1430
 Db 1636 GCTCATATTTGTTCTTACGCTTGGCCCGTGCATTATGAAGACCTTGAATTCGATTCGATTCATA 1695
 QY 1431 TTCCAGAATTTGAAGCTGTAAA 1451
 Db 1696 TGACAAATTCAGGCAGTAAA 1716

RESULT 37

US-09-620-312D-171
 ; Sequence 171, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Wang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yunding
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghaast
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pt_FL_genes Version 1.0
 ; SEQ ID NO 171
 ; LENGTH: 2809
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1936)..(2364)
 US-09-620-312D-171

Query Match 7.2%; Score 106.4; DB 4; Length 2809;
 Best Local Similarity 55.0%; Pred. No. 1.9e-23;
 Matches 231; Conservative 0; Mismatches 186; Indels 3; Gaps 1;
 QY 996 ACTAGGTACTGGCAGTATCAACCTCTACTCTAGTTCTTACTACAAACTATCTCA 1055
 Db 1851 ACTTGGCACTGGATAGCAGGCAATACCCTCAATCACTCATACCAACACTATTCAC 1910
 QY 1056 AGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTGCCTTGGCAAGATCAACT 1115
 Db 1911 AACCTTTCTAACACCGTAGAATATGCACACTTCCATTACCAGTCTCCAAAGCAAT 1970
 QY 1116 TAACCTCCCTAGCAGCAGTACTCTTCAAAATCGAAGCTTTAGACTTGTAAACCGCAA 1175
 Db 1971 AGACTTCTCTGGGGGTATCTCTTCAAACTGGAGATCCTGGACCTCTTAACCACTGA 2030
 QY 1176 AAGAGGGGGAACCTGTTTATTTTAGGAGAAGAACGCTGTTATATGTTAATCAATCCAG 1235

Db 2031 GAAAGGGGTACTTGCATATACCTCCAGGAAGAAATGCTGTTCTGTAAATGAATCTGG 2090
 QY 1236 AATTGTCACTAGAGAAAGTTAAAGAAATTCGAGATCG---AATACAATGTAGAGCAGAGGA 1292
 Db 2091 CATTTGTTTCATATCGCAGTTCTGTAGGCTTCATGACAGGGCTGCAGAGCTTAGACATCAAGT 2150
 QY 1293 GCTTCAAAACACCGACGCTGGGCGCTCTCAGCAATGGATGCCCTGGGTTCTTCCCTTT 1352
 Db 2151 CGCTGACTCTGTGGCAAGGATCATCCCTTCTAAGATGATACCTTGGTGGTGGCCCTT 2210
 QY 1353 CTTAGGACCTCTAGCAGCTCTAATATTGTTACTCTCTTTGGACCTGTATCTTTAACTT 1412
 Db 2211 CCTAGGACCCCTGATCTTCTCTTCTCTTACTAATGATTGGGCCATGATATTTAACTT 2270

RESULT 38

US-09-904-615-12
 ; Sequence 12, Application US/09904615
 ; Patent No. 6566325
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 49 Human Secreted Proteins
 ; FILE REFERENCE: P2032P1
 ; CURRENT APPLICATION NUMBER: US/09/904,615
 ; PRIOR FILING DATE: 2001-07-16
 ; PRIOR FILING DATE: 09/511,554
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR FILING DATE: 60/097,917
 ; PRIOR FILING DATE: 1998-08-25
 ; PRIOR FILING DATE: 60/098,634
 ; PRIOR FILING DATE: 1998-08-31
 ; NUMBER OF SEQ ID NOS: 170
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 12
 ; LENGTH: 2342
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-904-615-12

Query Match 7.1%; Score 104.8; DB 4; Length 2342;
 Best Local Similarity 54.8%; Pred. No. 5.7e-23;
 Matches 230; Conservative 0; Mismatches 187; Indels 3; Gaps 1;
 QY 996 ACTAGTACTGGCATTGGCAGTATCAAACTCTACTCAGTTCTTACTACAAACTATCTCA 1055
 Db 1357 ACTTGGCACTGGATAGCAGGCATAAACACCTCAATCACTCATACCAACACTATTTCAC 1416
 QY 1056 AGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTGCCTTGGCAAGATCAACT 1115
 Db 1417 AACCTTTCTAACACCGTAGAAGATATGCACACTTCCATTACCAGTCTCCAAACGCAAT 1476
 QY 1116 TAACCTCCCTAGCAGCAGTACTCTTCAAAATCGAAGACTTTAGACTTTGCTAAACGCAA 1175
 Db 1477 AGACTTCTCTGGGAGTCACTCTTCAAACTGGAGACTCTGGACCTCTTAACCACTGA 1536
 QY 1176 AAGAGGGGGAACCTGTTTATTTTAGGAGAAGAAAGCTGTTTATTTAGTTAATCAATCCAG 1235
 Db 1537 GAAAGGGGTACTCTGCATATACCTCCAGGAAGAAATGCTGTTCTGTGTTAATGAATCTGG 1596
 QY 1236 AATTGTCACTAGAGAAAGTTAAAGAAATTCGAGATCG---AATACAATGTAGAGCAGAGGA 1292
 Db 1597 CATTTGTTTCATATCGCAGTTCTGTAGGCTTCATGACAGGGCTGCAGAGCTTTGCATCAAGT 1656
 QY 1293 GCTTCAAAACACCGAACGCTGGGCGCTCTCAGCCAAATGGATGCCCTGGGTTCTCCCTTT 1352
 Db 1657 CGCTGACTCTCTGGTGGCAAGGATCATCCCTTCTAAGATGGATACCTGGTGGTGGCCCTT 1716
 QY 1353 CTTAGGACCTCTAGCAGCTCTAATAATTGTTACTCTCTTTGGACCTGTATCTTTAACTT 1412
 Db 1717 CCTAGGACCCCTGATCTTCTCTTCTCTTACTAATGATTGGGCCATGATATTTAACTT 1776

RESULT 39

```

US-09-949-016-12735/C
; Sequence 12735, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12735
; LENGTH: 80355
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12735

```

Search completed: February 21, 2005, 15:01:25
Job time : 267.948 secs

RESULT 40

```

RESULI 40
US-09-949-016-13572/c
; Sequence 13572, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 11:52:50 ; Search time 818.526 Seconds
(without alignments)
10694.149 Million cell updates/sec

Title: US-09-319-156B-9
Perfect score: 1481
Sequence: 1 atggccctcccttatcatc.....gtctacaagtgaacccca 1481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5384158 seqs, 2955248155 residues
Total number of hits satisfying chosen parameters: 10769316

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1481	100.0	1481	US-08-979-847-105	Sequence 105, App
2	1481	100.0	1481	US-10-114-104-105	Sequence 105, App
3	1481	100.0	2030	US-10-637-565-18	Sequence 18, Appl
4	1389.8	93.8	2074	US-10-416-642-4	Sequence 4, Appl
5	1335.4	90.2	2930	US-09-902-535-1	Sequence 1, Appl
6	1335.4	90.2	2946	US-10-114-893-134	Sequence 134, App
7	1335.4	90.2	2946	US-10-016-249-3	Sequence 3, Appl
8	1333.8	90.1	1617	US-10-133-036-3	Sequence 3, Appl
9	1333.8	90.1	8523	US-09-854-867-21	Sequence 21, Appl
10	1333.8	90.1	56093	US-09-873-367C-81	Sequence 81, Appl
11	1324.2	89.4	7582	US-10-632-793-30	Sequence 30, Appl

12	1308.2	88.3	2782	17	US-10-133-036-1	Sequence 1, Appl
13	1306.6	88.2	2782	17	US-10-632-793-26	Sequence 26, Appl
C 14	1277.4	86.3	161334	13	US-10-087-192-730	Sequence 730, App
15	1131.6	76.4	1894	9	US-09-864-761-4444	Sequence 4444, Ap
16	1116.4	75.4	1948	17	US-10-632-793-24	Sequence 24, Appl
17	716.6	48.4	792	9	US-09-864-761-21192	Sequence 21192, A
18	591.8	40.0	1684	17	US-10-363-616-228	Sequence 228, App
19	582.8	39.4	591	17	US-10-632-793-19	Sequence 19, Appl
20	516.6	34.9	1136	17	US-10-632-793-25	Sequence 25, Appl
21	513.6	34.7	2006	17	US-10-632-793-23	Sequence 23, Appl
C 22	377.8	25.5	570	13	US-10-027-632-322491	Sequence 322491, A
C 23	377.8	25.5	570	13	US-10-027-632-322492	Sequence 322492, A
C 24	377.8	25.5	570	17	US-10-027-632-322491	Sequence 322491, A
C 25	377.8	25.5	570	17	US-10-027-632-322492	Sequence 322492, A
C 26	364	24.6	521	16	US-10-029-386-4312	Sequence 4312, Ap
27	353	23.8	551	13	US-10-027-632-322574	Sequence 322574, A
28	353	23.8	551	13	US-10-027-632-322575	Sequence 322575, A
29	353	23.8	551	17	US-10-027-632-322574	Sequence 322574, A
30	353	23.8	551	17	US-10-027-632-322575	Sequence 322575, A
C 31	318.6	21.5	822900	17	US-10-292-798-1393	Sequence 1393, Ap
C 32	302	20.4	1329	8	US-08-979-847-108	Sequence 108, App
C 33	302	20.4	1329	16	US-10-114-104-108	Sequence 108, App
C 34	296	20.0	426	9	US-09-864-761-14030	Sequence 14030, A
C 35	229.8	15.5	625	13	US-10-027-632-316412	Sequence 316412, A
C 36	229.8	15.5	625	17	US-10-027-632-316412	Sequence 316412, A
C 37	229.6	15.5	619	13	US-10-027-632-86541	Sequence 86541, A
C 38	229.6	15.5	619	17	US-10-027-632-86541	Sequence 86541, A
C 39	219.4	14.8	559	9	US-09-864-761-7501	Sequence 7501, Ap
C 40	213	14.4	494	16	US-10-029-386-20259	Sequence 20259, A
C 41	194.8	13.2	520	9	US-09-864-761-7310	Sequence 7310, Ap
C 42	189.8	12.8	740	13	US-10-027-632-150080	Sequence 150080, A
C 43	189.8	12.8	740	17	US-10-027-632-150080	Sequence 150080, A
C 44	173.8	11.7	527	13	US-10-027-632-85218	Sequence 85218, A
C 45	173.8	11.7	527	17	US-10-027-632-85218	Sequence 85218, A

ALIGNMENTS

RESULT 1
US-08-979-847-105
; Sequence 105, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUREK, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC PURPOSES
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435

```

; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-979-847-105

Query Match      100.0%; Score 1481; DB 8; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTATCATACTTTTCTCTTTACTGTCTCTTACCCGCTTTCGCTCTCACT 60
DB 1 ATGGCCCTCCCTATCATACTTTTCTCTTTACTGTCTCTTACCCGCTTTCGCTCTCACT 60
QY 61 GCACCCCTCCATGCTGCTGTACACACGAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 61 GCACCCCTCCATGCTGCTGTACACACGAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
QY 121 AGCGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
DB 121 AGCGGCTTCTGGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGAAACTCC 180
QY 181 ACCTTCAGTCCGACACCATATGCCCCGCACTGCTATACTCGCCACTCTTTGCATG 240
DB 181 ACCTTCAGTCCGACACCATATGCCCCGCACTGCTATACTCGCCACTCTTTGCATG 240
QY 241 CATGCAAACTACTTATTGGACAGGAAATGATTAACTCTAGTTGCTCGAGGACTT 300
DB 241 CATGCAAACTACTTATTGGACAGGAAATGATTAACTCTAGTTGCTCGAGGACTT 300
QY 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCCATACAGTATGCTGATGGGGTGAATT 360
DB 301 GGAGCCACTGTCTGTTGGACTTACTTCAACCCATACAGTATGCTGATGGGGTGAATT 360
QY 361 CAAGGTGAGCAAGAGAAAACAGTAAAGGAGCACTCTCCCACTGACCCGGGACAT 420
DB 361 CAAGGTGAGCAAGAGAAAACAGTAAAGGAGCACTCTCCCACTGACCCGGGACAT 420
QY 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCCTCCGTACC 480
DB 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCCCTCCGTACC 480
QY 481 CATATCTCGGCTGTGAGCCCTATTAAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB 481 CATATCTCGGCTGTGAGCCCTATTAAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
QY 541 CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCATCTCAGGCCATACATTTCAATC 600
DB 541 CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCATCTCAGGCCATACATTTCAATC 600
QY 601 CTGTTTCTGAAACAATGGAACAATCTTACGACAGAAATAAACAACCACTTCCGTTTTAGTA 660
DB 601 CTGTTTCTGAAACAATGGAACAATCTTACGACAGAAATAAACAACCACTTCCGTTTTAGTA 660
QY 661 GGAACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCACTGTTGTAATAATTT 720
DB 661 GGAACCTCTGTTTCCAACTGGAATAAACCCATACCTCAAACTCACTGTTGTAATAATTT 720
QY 721 AGCAATACTATAGACACACCACTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
DB 721 AGCAATACTATAGACACACCACTCCCAATGATCAGGTGGGTAAACCTCCACACGA 780
QY 781 ATAGTCTGCGCTACCCCTCAGGAATATTTTTTGTCTGTGTGATCCTCAGCCCTATCATTTGTTG 840
```

```

DB 781 ATAGTCTGCGCTACCCCTCAGGAATATTTTTTGTCTGTGTGATCCTCAGCCCTATCATTTGTTG 840
QY 841 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTAGTGGCCCTATGACCATCTAC 900
DB 841 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTAGTGGCCCTATGACCATCTAC 900
QY 901 ACTGAACAAGATTTATACAATCATGTCTAAGCCCCACAAACAAAGAGTACCCATT 960
DB 901 ACTGAACAAGATTTATACAATCATGTCTAAGCCCCACAAACAAAGAGTACCCATT 960
QY 961 CTTCCTTTTGTATCAGAGCAGGAGTCTAGGACAGACTAGGTACTGGCATTGGCAGTATC 1020
DB 961 CTTCCTTTTGTATCAGAGCAGGAGTCTAGGACAGACTAGGTACTGGCATTGGCAGTATC 1020
QY 1021 ACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGTTGACATGAAACAG 1080
DB 1021 ACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAGAAATAAATGTTGACATGAAACAG 1080
QY 1081 GTCACTGACTCCCTGGTCACTTCAAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
DB 1081 GTCACTGACTCCCTGGTCACTTCAAGAGTCAACTTAACTCCCTAGCAGCAGTAGTCCTT 1140
QY 1141 CAAATCGAAGAGCTTTAGACTTGTAAACCCGCAAAAGAGGGGAACTGTTATTTT 1200
DB 1141 CAAATCGAAGAGCTTTAGACTTGTAAACCCGCAAAAGAGGGGAACTGTTATTTT 1200
QY 1201 GGAGAAAGACGCTGTTATTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
DB 1201 GGAGAAAGACGCTGTTATTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGAA 1260
QY 1261 ATTGAGATCGAATACAAATGTAGACAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 1320
DB 1261 ATTGAGATCGAATACAAATGTAGACAGAGAGCTTCAAAACACCGAACGCTGGGCGCTC 1320
QY 1321 CTCAGCAATGGATGCGCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 1321 CTCAGCAATGGATGCGCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
QY 1381 TTACTCTCTTTGACCCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAAT 1440
DB 1381 TTACTCTCTTTGACCCCTGTATCTTTAACTCTCTTAAAGTTTGTCTCTTCCAGAAAT 1440
QY 1441 GAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAACCCCA 1481
DB 1441 GAAGCTGTAAAGCTACAGATGCTCTTACAAATGGAACCCCA 1481
```

RESULT 2
US-10-114-104-105
; Sequence 105, Application US/10114104
; Publication No. US20030198647A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHY
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,104
FILING DATE: 03-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,847
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-10-114-104-105
Query Match 100.0%; Score 1481; DB 16; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCCCTCCCTATCATACATTTCTTTACTGTTCTTACCCCTTTCGCTCTCACT 60
DB 1 ATGGCCCTCCCTATCATACATTTCTTTACTGTTCTTACCCCTTTCGCTCTCACT 60
QY 61 GCACCCCTCCATGCTCTGACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
DB 61 GCACCCCTCCATGCTCTGACACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
QY 121 ACGGCGCTTCTGGAAATATTGATGCCCCCATCATATAGAGATTATCTAAGGAACTCC 180
DB 121 ACGGCGTTCTGGAAATATTGATGCCCCCATCATATAGAGATTATCTAAGGAACTCC 180
QY 181 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATAACTCTGTCACCTTTTGCATG 240
DB 181 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATAACTCTGTCACCTTTTGCATG 240
QY 241 CATGCAATATCTANTATTGGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 300
DB 241 CATGCAATATCTANTATTGGACAGGAAATGATTAATCTAGTTCTCTGGAGGACTT 300
QY 301 GGAGCCACTGCTGTGGACTTACTTCAACCCATACCAAGTAGTGTGATGGGGGTGAATT 360
DB 301 GGAGCCACTGCTGTGGACTTACTTCAACCCATACCAAGTAGTGTGATGGGGGTGAATT 360
QY 361 CAAGGTCAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGGACAT 420
DB 361 CAAGGTCAGGCAAGAGAAAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGGACAT 420
QY 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCAAACTACATGAAACCCCTCGTACC 480
DB 421 AGCACCCCTAGCCCTCAAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCGTACC 480
QY 481 CATACTGCGCTGGTGGAGCTATTTAATACCAACCCCTCACTCGGCTCCCATGAGGTCTCAGCC 540
DB 481 CATACTGCGCTGGTGGAGCTATTTAATACCAACCCCTCACTCGGCTCCCATGAGGTCTCAGCC 540
QY 541 CAAAACCCCTACTAACTGTGTGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 600
DB 541 CAAAACCCCTACTAACTGTGTGATGTGCTCCCTCGCACTTCAGGCCATACATTTCAATC 600

QY 601 CCTGTTCTGAAACAATGGAAACAATTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
DB CCTGTTCTGAAACAATGGAAACAATTTCAGCACAGAAATAAACACCACTTCCTGTTTAGTA 660
QY 661 GGACCTCTGTTTCCAAATCTGGAAATAACCCATACTCAAACTCCTACCTGTGTAAATTT 720
DB 661 GGACCTCTGTTTCCAAATCTGGAAATAACCCATACTCAAACTCCTACCTGTGTAAATTT 720
QY 721 AGCAATATCTATAGACACACACAGCTCCCAATCATCATGAGTGGGTAAACACTCCACACGA 780
DB 721 AGCAATATCTATAGACACACACAGCTCCCAATCATCATGAGTGGGTAAACACTCCACACGA 780
QY 781 ATAGTCTGCTACCCCTCAGGAATATTTTGTGTGGTACCTCAGCCCTATCATTTGTTG 840
DB 781 ATAGTCTGCTACCCCTCAGGAATATTTTGTGTGGTACCTCAGCCCTATCATTTGTTG 840
QY 841 AATGGCTCTTCAGAACTCTATGCTTCTCTCATTTCTTAGTGCCCCCTATGACCATCTAC 900
DB 841 AATGGCTCTTCAGAACTCTATGCTTCTCTCATTTCTTAGTGCCCCCTATGACCATCTAC 900
QY 901 ACTGAACAAGATTTATACAAATCATGTCGTACCTTAAGCCCCACAAACAAAGAGTACCATT 960
DB 901 ACTGAACAAGATTTATACAAATCATGTCGTACCTTAAGCCCCACAAACAAAGAGTACCATT 960
QY 961 CTTCTTTTGTATCAGACAGGAGTGTAGGACACTAGGTACTGGCATTTGCGAGTATC 1020
DB 961 CTTCTTTTGTATCAGACAGGAGTGTAGGACACTAGGTACTGGCATTTGCGAGTATC 1020
QY 1021 ACNACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGTCGACATGGAACAG 1080
DB 1021 ACNACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGTCGACATGGAACAG 1080
QY 1081 GTCACTGACTCCCTGCTGCTTGCAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT 1140
DB 1081 GTCACTGACTCCCTGCTGCTTGCAGATCAACTTTAACTCCCTAGCAGCAGTAGTCCTT 1140
QY 1141 CAAATCGAAGAGCTTTAGACTTTGCTAACCGCCAAAGAGGGGGAACCTGTTTATTTTA 1200
DB 1141 CAAATCGAAGAGCTTTAGACTTTGCTAACCGCCAAAGAGGGGGAACCTGTTTATTTTA 1200
QY 1201 GGAAGAAGCGCTTATTATGTTAATCAATCAATCAGAAATGCTCACTGAGAAGTTAAAGAA 1260
DB 1201 GGAAGAAGCGCTTATTATGTTAATCAATCAATCAGAAATGCTCACTGAGAAGTTAAAGAA 1260
QY 1261 ATTGAGATCGAATACAATGTCAGACAGAGAGCTTCAAAACACCGAACCTCGGGCCTC 1320
DB 1261 ATTGAGATCGAATACAATGTCAGACAGAGAGCTTCAAAACACCGAACCTCGGGCCTC 1320
QY 1321 CTCAGCAATGGATGCGCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB 1321 CTCAGCAATGGATGCGCTGGGTTCTCCCTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
QY 1381 TTACTCTCTTTGGACCCCTGATCTTTAACTCCTCTGTTAAGTTTGTCTCTTCCAGAAAT 1440
DB 1381 TTACTCTCTTTGGACCCCTGATCTTTAACTCCTCTGTTAAGTTTGTCTCTTCCAGAAAT 1440
QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
DB 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481

RESULT 3

US-10-637-565-18
; Sequence 18, Application US/10637565
; Publication No. US20040043381A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: PERRON, Hervé
; APPLICANT: KOMURIAN-PRADEL, Florence
; TITLE OF INVENTION: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES
; FILE REFERENCE: 110257
; CURRENT APPLICATION NUMBER: US/10/637,565

```
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US/09/890,340
; PRIOR FILING DATE: 2003-11-27
; PRIOR APPLICATION NUMBER: PCT/IB00/00159
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: EP 99420041.8
; PRIOR FILING DATE: 1999-02-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 2030
; TYPE: DNA
; ORGANISM: MSRV-1 retrovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1626)
; US-10-637-565-18

Query Match      100.0%; Score 1481; DB 17; Length 2030;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60
DB      1  ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60

QY     61  GCACCCCTCCATGCTGCTGTACACACAGTAGTCCCTTACCAAGAGTTTCTATGAGA 120
DB     61  GCACCCCTCCATGCTGCTGTACACACAGTAGTCCCTTACCAAGAGTTTCTATGAGA 120

QY     121  AGCGGCTTCTCGGAATATTGATGCCCATATATAGGAGTTTATCTAAGGGAATCTCC 180
DB     121  AGCGGCTTCTCGGAATATTGATGCCCATATATAGGAGTTTATCTAAGGGAATCTCC 180

QY     181  ACCTTCACGCCACACACCATATGCCCGCAACTGCTATACTTGCCACTCTTTGCATG 240
DB     181  ACCTTCACGCCACACACCATATGCCCGCAACTGCTATACTTGCCACTCTTTGCATG 240

QY     241  CATGCAATACTCATTTATGGACAGGGAATGATTAATCTAGTTGCTCGGAGACTT 300
DB     241  CATGCAATACTCATTTATGGACAGGGAATGATTAATCTAGTTGCTCGGAGACTT 300

QY     301  GGAGCCACTGCTGTTGGACTTACTTACCCCATACAGTATGCTGTGATGGGGTGGAAAT 360
DB     301  GGAGCCACTGCTGTTGGACTTACTTACCCCATACAGTATGCTGTGATGGGGTGGAAAT 360

QY     361  CAAGGTCAGGCAAGAGAAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT 420
DB     361  CAAGGTCAGGCAAGAGAAAAACAAGTAAAGAGCAATCTCCAACTGACCCGGGACAT 420

QY     421  AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAACTACATGAAACCTCCGTACC 480
DB     421  AGCACCCCTAGCCCTTACAAAGACTAGTTCTCTCAAACTACATGAAACCTCCGTACC 480

QY     481  CATACTCGGCTGGTGAAGCTATTAAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540
DB     481  CATACTCGGCTGGTGAAGCTATTAAATACACCCCTCACTCGGCTCCATGAGTCTCAGCC 540

QY     541  CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAAGGCAATCAATC 600
DB     541  CAAAACCCCTACTAACTGTTGGATGTGCTCCCTCCCTGCACCTTCAAGGCAATCAATC 600

QY     601  CTTGTTCTTGAACATGGAACAACCTTACGACAGAAATAAACACCACTTCCGTTTAGTA 660
DB     601  CTTGTTCTTGAACATGGAACAACCTTACGACAGAAATAAACACCACTTCCGTTTAGTA 660

QY     661  GGACCTCTTGTGTTTCCAAATCTGGAATAAACCACTCACTCACTGTGTAAATTT 720
DB     661  GGACCTCTTGTGTTTCCAAATCTGGAATAAACCACTCACTCACTGTGTAAATTT 720

QY     721  AGCAATACTATAGACACACAGCTCCCAATGCATGAGTGGGTAACTCCCAACAGA 780
DB     721  AGCAATACTATAGACACACAGCTCCCAATGCATGAGTGGGTAACTCCCAACAGA 780
```

RESULT 4

```
US-10-416-642-4
; Sequence 4, Application US/10416642
; Publication No. US20040043452A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: ARVIZU, Chandra
; TITLE OF INVENTION: EMBRYOGENESIS ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0842 PCT
; CURRENT APPLICATION NUMBER: US/10/416,642
; CURRENT FILING DATE: 2003-05-13
; PRIOR APPLICATION NUMBER: 60/249,407
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 2074
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040043452A1 7477736CB1
```

```
781  ATAGTCTGCTACCCCTCAGGAATATTTTCTGTGTGTACCTCAGCCTATCATTTGTTG 840
DB   781  ATAGTCTGCTACCCCTCAGGAATATTTTCTGTGTGTACCTCAGCCTATCATTTGTTG 840

QY   841  AATGGCTCTTCAGAACTTATGCTTCTCTCATTTTAGTGCCCTTATGACCATCTAC 900
DB   841  AATGGCTCTTCAGAACTTATGCTTCTCTCATTTTAGTGCCCTTATGACCATCTAC 900

QY   901  ACTGAAACAAGATTTATACAATCATGTCTAGTACGCCCCCACAAAGAGTAGCCATT 960
DB   901  ACTGAAACAAGATTTATACAATCATGTCTAGTACGCCCCCACAAAGAGTAGCCATT 960

QY   961  CTTCTCTTTTGTATTCAGAGCAGGAGTGTAGGACAGACTAGGTACTGGCAATGGCAGTATC 1020
DB   961  CTTCTCTTTTGTATTCAGAGCAGGAGTGTAGGACAGACTAGGTACTGGCAATGGCAGTATC 1020

QY   1021  ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080
DB   1021  ACAACCTCTACTCAGTTCTTACTACAACTATCTCAAGAAATAAATGGTGACATGGAACAG 1080

QY   1081  GTCACTGACTCCCTGGTCACTTGCAGAGTCAACTTAACCTTAACTCCCTAGCAGCAGTAGTCCTT 1140
DB   1081  GTCACTGACTCCCTGGTCACTTGCAGAGTCAACTTAACCTTAACTCCCTAGCAGCAGTAGTCCTT 1140

QY   1141  CAAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAGAGGGGAACTGTTTATTTTAA 1200
DB   1141  CAAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAGAGGGGAACTGTTTATTTTAA 1200

QY   1201  GGAGAGAAACGCTGTTATTTATTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA 1260
DB   1201  GGAGAGAAACGCTGTTATTTATTAATCAATCCAGAAATGTCACTGAGAAAGTTAAAGAA 1260

QY   1261  ATTGAGATCGAATACATGTTAGACAGAGAGCTTCAAAACACGACGCTGGGGCCTC 1320
DB   1261  ATTGAGATCGAATACATGTTAGACAGAGAGCTTCAAAACACGACGCTGGGGCCTC 1320

QY   1321  CTCAGCCAAATGGATGCCCTGCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
DB   1321  CTCAGCCAAATGGATGCCCTGCTCCCTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380

QY   1381  TTACTCTCTTTGGACCCCTGATCTTTAACCTCTGTTAAAGTTTGTCTCTTCCAGAAAT 1440
DB   1381  TTACTCTCTTTGGACCCCTGATCTTTAACCTCTGTTAAAGTTTGTCTCTTCCAGAAAT 1440

QY   1441  GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA 1481
DB   1441  GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAAACCCCA 1481
```

```

;
; NAME/KEY: unsure
; LOCATION: 1995
; OTHER INFORMATION: a, t, c, g, or other
; US-10-416-642-4

Query Match      93.8%; Score 1389.8; DB 17; Length 2074;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1424; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 1 ATGCCCTCCCTTATCATCTTTCTCTTACTGTTCTTACCCCTTTCGCTCTCACT 60
Db 64 ATGCCCTCCCTTATTTCTCTTACTGTTCTCTCACCACCTTTCAGTCTCACT 123

Qy 61 GCACCCCTCCATGCTCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 124 GCACCCCTCTCATGCCACTGTAGACCAAGTAGCTCCCTTACCAGGGTTTCTATGGAGA 183

Qy 121 ACGCGGCTTCTCGAATAATTTGATGCCCCCATCATATATAGGAGTTTATCTAAGGGAAACTCC 180
Db 184 ATGGCGGTCCAGACATATTTGATGCCCCCATCGTATAGGAGTTTATCTAAGGGAAACCCC 243

Qy 181 ACCTTCACTGCCACACCCATATGCCCCGCAACCTGTATATCTGCTGCACTCTTTGCAATG 240
Db 244 GCCTTCAACGCCACACCCATATGCCCACTGTATTAACCTGCACTCTTTGTAATG 303

Qy 241 CATGCAATCTCATTTATGACAGGGAATGATTAATCTAGTTGCTGCTGGAGGACTT 300
Db 304 CATGCAATCTCATTTATGACAGGGAATGATTAATCTAGTTGCTGCTGGAGGACTT 363

Qy 301 GGAGCACTGTCTGTGCACTTACTTCAACCCATACCAAGTATGCTGATGGGGTGGAAAT 360
Db 364 GGAGCACTGTCTGTGCACTTACTTCAACCCATACCGGTATGCTGATGGGGTGGAGTT 423

Qy 361 CAAGGTGAGGCAAGAGAAAACAAGTAAAGGAAGCAATCTCCAACTGACCCGGGGACAT 420
Db 424 CAAGATCAGGCAAGAGAAAACAATGAAAGGAAGTAACTCTCCAACTCACCCGGGTACAT 483

Qy 421 AGCACCCCTAGCCCTCAAGAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db 484 AGCACCCCTAGCCCTCAAGAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC 543

Qy 481 CATACTCGCTGTGTGAGCTTATTAATACCAACCCCTCACTGGCTCCATGAGGTCTCAGCC 540
Db 544 CATACTCACCTGTGTGAGCTTATTAATACCAACCCCTCACTGGCTCCATGAGGTCTCAGCC 603

Qy 541 CAAACCCCTACTAACTGTGTGAGTGTCTCCCTGCACTTTCAGGCAATACATTTCAATC 600
Db 604 CAAACCCCTACTAACTGTGTGAGTGTCTCCCTGCACTTTCAGGCAATACATTTCAATC 663

Qy 601 CCGTGTCTGCAATGGAACAACTTCAGCAGAGAAATTAACCACTTCCGTTTGTAGTA 660
Db 664 CCGTGTCTGCAATGGAATTAATCTCAGCAGAGAAATTAACCACTTCCATTTTGTAGTA 723

Qy 661 GGACCTCTGTGTTTCCAACTCGAAATAACCCATACCTCAAACTTCACTGTGTAAATTT 720
Db 724 GGACCTCTGTGTTTCCAACTCGAAATAATCCATACCTCAAACTTCACTGTGTAAATTT 783

Qy 721 AGCAATACTATAGACACACAGCTCCCAATGCAATCAGGTGGGTAAACCTCCACACGA 780
Db 784 AGCAATACTATAGACACACCAACTCCCAATGCAATCAGGTGGGTAAACCTCCACACGA 843

Qy 781 ATAGTCTGCTTACCTCAGGAATATTTTGTGCTGTGTGATCTCAGGCTATCATTTGTTG 840
Db 844 ATAGTCTGCTTACCTCAGGAATATTTTGTGCTGTGTGATCTCAGGCTATCATTTGTTG 903

Qy 841 AATGCTCTTCCAGAACTATGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
Db 904 AATGCTCTTCCAGAACTATGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 963

Qy 901 ACTGAACAGATTTTATCAATCATGTGCTAGCTAAGCCCAACAACAAAGAGTACCCATT 960
Db 964 ACTGAACAGATTTTATCAATCATGTGCTAGCTAAGCCCAACAACAAAGAGTACCCATT 1023

```

RESULT 5

```

US-09-902-535-1
; Sequence 1, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: Mi, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preeclampsia and gestational trophoblast
; TITLE OF INVENTION: disorders
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930) ... (2546)
US-09-902-535-1

```

```

Query Match      90.2%; Score 1335.4; DB 9; Length 2930;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 1 ATGCCCTCCCTTATCATCTTTCTCTTACTGTTCTTACCCCTTTCGCTCTCACT 60
Db 930 ATGCCCTCCCTTATCATATTTTCTCTTACTGTTCTTTTACCCCTTTCCTCTCACT 989

```

QY	61	GCACCCCTCCATGCTGCTGTATCAACACGATAGCTCCCTTACCAAGAGTTTCTATGAAGA	120
DB	990	GCACCCCTCCATGCGCGCTGTATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGGAGA	1049
QY	121	ACGCGGCTTCTGTGGAATATTGATGCCCCATCATATAGGAGTTTATCTAAGGGAACCTCC	180
DB	1050	ATGCAGGTCCTCCGGAATATTGATGCCCATCGTATAGAGTCTTCTTAAGGGAACCCCCC	1109
QY	181	ACCTTCACCTGCCACACCCATATGCCCGGAACTGCTATAACTCTGCGCACTCTTTGCAATG	240
DB	1110	ACCTTCACCTGCCACACCCATATGCCCGGCAACTGCTATCACTCTGCCACTCTTTGCAATG	1169
QY	241	CATGCAAACTACTCAATTATTGGACAGGGAATAATTAACTCTAGTTGTCTCTGGAGACTT	300
DB	1170	CATGCAAACTACTCAATTATTGGACAGGGAATAATTAACTCTAGTTGTCTCTGGAGACTT	1229
QY	301	GGAGCCACTCTCTGTTGGACTTACTTCACCCATACCACTATGCTGTGATGGGGTGGAAATT	360
DB	1230	GGAGTCACTGCTCTGTTGGACTTACTTCACCCAACTGGTATGCTGTATGGGGTGGAGTT	1288
QY	361	CAAGGTCAGCAAGAGAAAACAAGTAAGGAAGCAATCTCCCACTGACCCCGGGACAT	420
DB	1290	CAAGATCAGCAAGAGAAAACAATGTAAGAAGTAACTCTCCCACTGACCCCGGGTACAT	1349
QY	421	AGCACCCCTAGCCCTACAAAGACTAGTTCTCTCAAAACTCATGAACCCCTCCGTACC	480
DB	1350	GGCACTCTAGCCCTACAAAGACTAGATCTCTCAAAACTCATGAACCCCTCCGTACC	1409
QY	481	CATACCTGCGCTGGTGAGCCTATTAAATAACACCTCACTGGGTCCATGAGGTCTCAGCC	540
DB	1410	CATACCTGCGCTGGTAAGCCTATTAAATACCACTCACTGGGTCCATGAGGTCTCGGC	1469
QY	541	CAAAACCCCTACTAATCTGTTGGATGTCCTCCCTGCATCTTAGGSCCATACATTTCAATC	600
DB	1470	CAAAACCCCTACTAATCTGTTGGATGTCCTCCCTGAACTCTAGGSCCATATGTTTCAATC	1529
QY	601	CTCTGTCCTGAAACAATGGAAACAACCTCAGCACAGAAATAAACCACTTCCTGTTTAGTA	660
DB	1530	CCTGTACTGTAAACAATGGAAACAACCTCAGCACAGAAATAAACCACTTCCTGTTTAGTA	1589
QY	661	GGACCTCTTGTTTCCAACTCTGGAAATAACCCATACCTCAAACTCACTGTGTAAATTTT	720
DB	1590	GGACCTCTTGTTTCCAACTCTGGAAATAACCCATACCTCAAACTCACTGTGTAAATTTT	1649
QY	721	AGCAATACTATAGACACAACACGCTCCCAATGCATCAGTGGGTAAACACCTCCCAACAGA	780
DB	1650	AGCAATACTATACATACAAACCACTCCCAATGCATCAGTGGGTAACTCTCCCAACAGA	1709
QY	781	ATAGTCTGCTTACCTCAGGAATAATTTTTGTCTGTGTACCTCAGCCTATCACTGTTG	840
DB	1710	ATAGTCTGCTTACCTCAGGAATAATTTTTGTCTGTGTACCTCAGCCTATCACTGTTG	1769
QY	841	AATGGCTCTTCAGAACTATATGTGCTTCTCTCATTTCTAGTGCCCCCTATGACCACTAC	900
DB	1770	AATGGCTCTTCAGAACTATATGTGCTTCTCTCATTTCTAGTGCCCCCTATGACCACTAC	1829
QY	901	ACTGAAACAGATTTTATACAATCATGTGCTACTAAGCCCCCAACAACAAAGAGTACCCATT	960
DB	1830	ACTGAAACAGATTTTATACAATCATGTGCTACTAAGCCCCCAACAACAAAGAGTACCCATT	1889
QY	961	CTTCTCTTTTGTATCAGACAGGAGTGTAGGCACTAGGTACTGTCATCTGGCGGTATC	1020
DB	1890	CTTCTCTTTTGTATAGGAGCAGGAGTGTAGGTGCACTAGGTACTGTCATCTGGCGGTATC	1949
QY	1021	ACAACCTCTACTCAGTTCTACTACAACTATCTCAGNAATAAATGTTGATCAGTACAG	1080
DB	1950	ACAACCTCTACTCAGTTCTACTACAACTATCTCAGNAATAAATGTTGATCAGTACAG	2009
QY	1081	GTCACCTGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGAGTAGTCTCTT	1140
DB	2010	GTCGCGGACTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGAGTAGTCTCTT	2069
QY	1141	CAAAATCGAAGAGCTTTTATGCTGTAAACCCGCAAAAGAGGGGAAACCTGTTTATTTT	1200

Db	2070	CGAAGAGCTTTAGACTGCTTAAACCGCTGAAAGAGGGGAAACCTGTTATTTTAA	2129
Qy	1201	GGAGAAAGACGCTGTTATTTATGTTAAATCAATCAGAAATTTGTCACCTGAGAAAGTTAAAGAA	1260
Db	2130	GGGGAAGAATGCTGTTATTTATGTTAAATCAATCGGAATCGTCACTGAGAAAGTTAAAGAA	2189
Qy	1261	ATTGAGATCGAATACAATGTTAGAGCAGAGGAGCTTCAAAACACCGAAACGCTGGGCGCTC	1320
Db	2190	ATTGAGATCGAATACAACGTTAGAGCAGAGGAGCTTCAAAACACCGAAACGCTGGGCGCTC	2249
Qy	1321	CTCAGCGAATGGATGCCCTGGGTTCTCCCTCTCTTAGGACCTCTAGCAGCTCTAAATATTG	1380
Db	2250	CTCAGCGAATGGATGCCCTGGATCTCCCTCTCTTAGGACCTCTAGCAGCTAAATATTG	2309
Qy	1381	TTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCCTTGTGTTAAAGTTGTTCTTCCAGATT	1440
Db	2310	CTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCCTTGTGTTAACTTTGTTCTTCCAGATC	2369
Qy	1441	GAGCTGTAAGCTACAGATGGTCTTACAATGGAACCCCA	1481
Db	2370	GAACTGTAAACTTACAAATGGAGCCCAAGATGCAATGCCAA	2410
RESULT 6			
US-10-114-893-134			
; Sequence 134, Application US/10114893			
; Publication No. US20020193567A1			
; GENERAL INFORMATION:			
; APPLICANT: McCoy, Kenneth			
; APPLICANT: McCoy, John M.			
; APPLICANT: LaVallie, Edward R.			
; APPLICANT: Collins-Racie, Lisa A.			
; APPLICANT: Evans, Cheryl			
; APPLICANT: Merberg, David			
; APPLICANT: Treacy, Maurice			
; APPLICANT: Bowman, Michael R.			
; APPLICANT: Spaulding, Vikki			
; APPLICANT: Carlin-Duckett, McKeough			
; APPLICANT: Kelleher, Kerry S.			
; APPLICANT: Genesys Institute, Inc.			
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM			
; FILE REFERENCE: GI 6000-10A			
; CURRENT APPLICATION NUMBER: US/10/114,893			
; CURRENT FILING DATE: 2002-04-02			
; EARLIER APPLICATION NUMBER: 09/413,232			
; EARLIER FILING DATE: 1999-10-06			
; NUMBER OF SEQ ID NOS: 321			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 134			
; LENGTH: 2946			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-114-893-134			
Query Match 90.2%; Score 1335.4; DB 13; Length 2946;			
Best Local Similarity 93.9%; Pred. No. 0;			
Matches 1390; Conservative 0; Mismatches 91; Indels 0; Gaps 0			
Qy	1	ATGGCCCTCCCTTATCATATCTTTCTTTTACTGTTCTTCTTACCCCTTTCGCTCTCACT	60
Db	928	ATGGCCCTCCCTTATCATATCTTTCTTTTACTGTTCTTTTACGCTCTTTCACTCTCACT	987
Qy	61	GCACCCCTCCATGCTGTGTAACACGAGTAGTCCCTTACCAGAGTTTCTATGAGA	120
Db	988	GCACCCCTCCATGCTGTGTAACACGAGTAGTCCCTTACCAGAGTTTCTATGAGA	104
Qy	121	ACGCGGCTTCTCGAAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAACCTC	180
Db	1048	ATGCAGCTCCCGAAATATTGATGCCCCCATCGTATAGGAGTTTCTAAGGGAACCTC	1107
Qy	181	ACCTTCACTGGCCACACCCATATGCCCGCACTGCTATACCTGTCACCTTTTGCATG	240

Qy	421	AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAAACTCATGAAACCCCTCCGTACC	480
Db	1348	GGCACTCTAGCCCTACAAAGGACTAGATCTCTCAAAACTCATGAAACCCCTCCGTACC	1407
Qy	481	CATACTCGCCTGGTGAGCCTATTATATACACACCTCACTCGGCTCCATGAGGTCTCAGCC	540
Db	1408	CATACTCGCCTGGTAGCCCTATTATATACACCCCTCACTGGGCTCCATGAGGTCTCGGC	1467
Qy	541	CAAAACCCCTACTAATCTTGGATGTGCTCCCTCCCTGCACCTTCAAGGCATACATTTCAATC	600
Db	1468	CAAAACCCCTACTAATCTTGGATGTGCTCCCTCCCTGAACTTCAAGGCATATGTTCAATC	1527
Qy	601	CTGTGCTGAAACAAATGGAACAACCTTCAGACAGAGAAATAAACACCACCTTCGGTTTATGTA	660
Db	1528	CTGTGCTGAAACAAATGGAACAACCTTCAGACAGAGAAATAAACACCACCTTCGGTTTATGTA	1587
Qy	661	GGACCTCTGTGTTTCCAACTCTGGAAATAACCCATACCTCAAACTCTGCTGTGAAATTTT	720
Db	1588	GGACCTCTGTGTTTCCAACTCTGGAAATAACCCATACCTCAAACTCTGCTGTGAAATTTT	1647
Qy	721	AGCAATACTATAGACACAACACGCTCCCAATGCAATGAGTGGGTAAACACCTCCACACGA	780
Db	1648	AGCAATACTATACACAACAACCTCCCAATGCAATGAGTGGGTAACTCCTCCACACAA	1707
Qy	781	ATAGTCTGCCCTACCCCTCAGGAATATTTTTGCTGTGTGTCCTCAGCCTATCATTTGTTG	840
Db	1708	ATAGTCTGCCCTACCCCTCAGGAATATTTTTGCTGTGTGTCCTCAGCCTATCGTTGTTG	1767
Qy	841	AATGGCTCTTCAGAAATCTATGTGCTCTCTCTCATTTCTAGTGCCCCCTATGACCACTAC	900
Db	1768	AATGGCTCTTCAGAAATCTATGTGCTCTCTCTCATTTCTAGTGCCCCCTATGACCACTAC	1827
Qy	901	ACTGAAACAAGATTTATACAATCATGTGCTACCTTAAGCCCCACACAACAAAGTACCCATT	960
Db	1828	ACTGAAACAAGATTTATACAATTTATGTCTATCTAAGCCCCGCGAACAAAAGAGTACCCATT	1887
Qy	961	CTTCTCTTTGTTTATCAGACAGGAGTGTCTAGGCGACAGCTAGTGTCTGGCATTTGGCGGTATC	1020
Db	1888	CTTCTCTTTGTTTATAGGACGAGGAGTGTCTAGGTGCACTAGTGTCTGGCATTTGGCGGTATC	1947
Qy	1021	ACAACTCTACTCAGTTCTACTACAACATATCTCAAGAAATAATGCTGATCATGGAACAG	1080
Db	1948	ACAACTCTACTCAGTTCTACTACAACATATCTCAAGAACTAATATGGGACATGGAACGG	2007
Qy	1081	GTCACTGACTCCCTGGTCACCTTCAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTT	1140
Db	2008	GTGCGGACTCCCTGGTCACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTT	2067
Qy	1141	CAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAAGAGGGGAACTGTGTTATTTTTA	1200
Db	2068	CAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAAGAGGGGAACTGTGTTATTTTTA	2127
Qy	1201	GGAGNAGAACGCTGTATTATGTTAATCAANTCAGATTTGTCCTGAGAAAGTTAAAGAA	1260
Db	2128	GGGGAAGAATCTGTATTATTTAATCAANTCCGGAATCGTCTGAGAAAGTTAAAGAA	2187
Qy	1261	ATTCGAGATCGAATACAATGTAGACGAGGAGCTTCAAAACACCGAACCGCTGGGGCTC	1320
Db	2188	ATTCGAGATCGAATACAACGTAGACGAGGAGCTTCGAACACTGGACCCCTGGGGCTC	2247
Qy	1321	CTCAGCCAAATGGATGCCCTCGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG	1380
Db	2248	CTCAGCCAAATGGATGCCCTCGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTATTAATATTG	2307
Qy	1381	TTACTCTCTTTGGACCCGTGTATCTTTAACTCTCTTGTAAAGTTGTCTCTTCCAGAAAT	1440
Db	2308	CTACTCTCTTTGGACCCGTGTATCTTTAACTCTCTTGTAACTTTGTCTCTTCCAGATC	2367
Qy	1441	GAAGCTGTAAAGCTACAGATGGTCTTACAATATGGAACCCCA	1481
Db	2368	GAAGCTGTAAAGCTACAATATGGAGCCCAAGATGCACTCAA	2408

```

RESULT 8
US-10-133-036-3
; Sequence 3, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
;
; APPLICANT: Conrad, Bernard
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Human endogenous retrovirus
US-10-133-036-3

```

Query Match	90.1%	Score 1333.8	DB 17	Length 1617
Best Local Similarity	93.8%	Pred. No. 0		
Matches 1389	Conservative 0	Mismatches 92	Indels 0	Gaps 0
QY	1	ATGSCCCTCCCTTATCATCTTTTCTCTTACTGTCTCTTACCCCTTTTGCCTCTCACT	60	
DB	1	ATGSCCCTCCCTTATCATATTTTCTTACTGTCTCTTACCTTTTACCTCTTTCACCTCACT	60	
QY	61	GCACCCCTCCATGCTGCTGTACAAACAGTAGTCCCTTTACCAAGAGTTTTCTATGAAGA	120	
DB	61	GCACCCCTCCATGSCCGCTGTATGACCAAGTAGCTCCCTTACCAAGAGTTTTCTATGGAGA	120	
QY	121	ACGGGCTTCTGGAATATTTGATGCCCCATCATATAGGAGTTTTCTAAGGGAAACCTCC	180	
DB	121	ATGCAGGCTCCCGAAATATTTGATGCCCCATCGTATAGGAGTCTTTCTAAGGGAAACCCCC	180	
QY	181	ACCTTCACTGCCACACCATATGCCCCGCAACTGCTATAACTCTGCCACTCTTTTGCATG	240	
DB	181	ACCTTCACTGCCCAACCATATGCCCCGCAACTGCTATCACTCTGCCACTCTTTTGCATG	240	
QY	241	CATGCAAACTACTCATTTATGGAACAGGGAAATGATTAATCTAGTTGTCTTGGAGGACTT	300	
DB	241	CATGCAAACTACTCATTTATGGAACAGGGAAATGATTAATCTAGTTGTCTTGGAGGACTT	300	
QY	301	GGAGCCACTGTCTGTTGGACTTACTTACCCCATACCAAGTAGTCTGATGGGGGTGGAATT	360	
DB	301	GGAGTCATGTCTGTTGAGCTTACTTACCCAAACTGATATGTCTGATGGGGGTGGAGTT	360	
QY	361	CAAGGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTCACCCCGGGACAT	420	
DB	361	CAAGATCAGGCAGAGAAAAACATGTAAGAAGTATCTCCCAACTCACCCGGGTACAT	420	
QY	421	AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC	480	
DB	421	GGCACCTTAGCCCTTACAAAGGACTAGATCTCTCAAAACTACATGAAACCCCTCCGTACC	480	
QY	481	CATACTGCGCTGTGAGCGCTATTTAATACACCCCTCACTCGGTCCATGAGTCTCAGCC	540	
DB	481	CATACTGCGCTGGTAAGCCTATTTAATACCAACCCCTCACTGGGTCCATGAGTCTCGGCC	540	
QY	541	CAAAACCCCTACTACTGTGAGTGTGCTCCCTCGCACTTTCAGGCCATACATTTTCAATC	600	
DB	541	CAAAACCCCTACTACTGTGAGTATGCTCCCTCGCACTTTCAGGCCATATGTTTCAATC	600	
QY	601	CCTGTTCTCTGAAACATGAAACAACCTTTCAGCAGAGAAATAAACCACTTCGGTTTAGTA	660	
DB	601	CCTGTACTCTGAAACATGAAACAACCTTTCAGCAGAGAAATAAACCACTTCGGTTTAGTA	660	
QY	661	GGACCTCTGTTTTCAAATCTCGGAATAACCCNTACTCAACCTCAACCTGTGTAATTT	720	
DB	661	GGACCTCTGTTTTCAAATCTCGGAATAACCCNTACTCAACCTCAACCTGTGTAATTT	720	

Db 7809 CTTCTTTTGTATAGGACGAGTGTAGTGCATCTAGTACTGGCATTTGGCGGTATC 7868
Qy 1021 ACAACCTCTACAGTTCTACTACAACTATCTCAAGAAATAATATGGTGAACATGGAACAG 1080
Db 7869 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGG 7928
Qy 1081 GTCACTGACTCCTGGTGCACCTTGCAGATCACTTAACCTCCTAGCAGCAGTACTCCTT 1140
Db 7929 GTGCGCGACTCCTGGTGCACCTTGCAGATCACTTAACCTCCTAGCAGCAGTACTCCTT 7988
Qy 1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGGAACCTGTTTATTTTA 1200
Db 7989 CAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGGAACCTGTTTATTTTA 8048
Qy 1201 GGAGAGAAACGCTGTTTATTTATTTATCAATCAATCCAGAAATGTCTACAGAGAGTTAAAGAA 1260
Db 8049 CGGGAAGAAATGCTGTTTATTTATTTATCAATCCGGAATCGTCACTGAGAGAGTTAAAGAA 8108
Qy 1261 ATTGAGATCGAATCAATGTAGACGAGAGGACTTCAAAACACCGAAGCTGGGGCCTC 1320
Db 8109 ATTGAGATCGAATCAACGTAAGACGAGAGGACTTCAAAACACCTGGACCTGGGGCCTC 8168
Qy 1321 CTCAGCAATGGATGGATCCCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTG 1380
Db 8169 CTCAGCAATGGATGGATCCCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTAATATTG 8228
Qy 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTTTTAACTTTGTCTCTCCAGAAAT 1440
Db 8229 CTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTTTTAACTTTGTCTCTTTCCAGATC 8288
Qy 1441 GAAGCTGTAAGCTACAGATGCTTTACAAATGGAGCCCAAGATGCGATCCAA 1481
Db 8289 GAAGCTGTAAGCTACAAATGGAGCCCAAGATGCGATCCAA 8329

RESULT 10

US-09-873-367C-81
; Sequence 81, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-81

Query Match 90.1%; Score 1333.8; DB 10; Length 56093;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATACTTTTCTCTTTACTGTTCTTTTACCCCTTTGGCTCTCACT 60

Db 35879 ATGGCCCTCCCTTATCATATTTTCTTTACTGTTCTTTTATACCTCTCTTTCACCTCTCACT 35938
Qy 61 GCACCCCTCCATGCTCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTTATGAAGA 120
Db 35939 GCACCCCTCCATGCCCTGTATGACGAGTAGTCCCTTACCAAGAGTTTCTTATGGAGA 35998
Qy 121 ACGGGCTTCTCTGGAATATTTGATGCCCATCATATAGGAGTTTATCTAAGGGAAACTCC 180
Db 35999 ATGACGGTCCCGGAATATTTGATGCCCATCTGTATAGGAGTCTTTCTAAGGGAAACCCC 36058
Qy 181 ACCTTCTCTGCCACACCCATATGCCCCGGCACTGTCTAATCTGTCACCTCTTTTGCATG 240
Db 36059 ACCTTCTCTGCCACACCCATATGCCCCGGCACTGTCTAATCTGTCACCTCTTTTGCATG 36118
Qy 241 CATGCAATCTCTATTTTGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT 300
Db 36119 CATGCAATCTCTATTTTGGACAGGAAATGATTAATCTAGTTGCTCTGGAGGACTT 36178
Qy 301 GGAGCCACTGCTCTGTTGACTTACTTCACCCATACCACTATGTCTGATGGGGTGGAAAT 360
Db 36179 GGAGTCACTGCTCTGTTGACTTACTTCACCCAATCTGTATGTCTGATGGGGTGGAGTT 36238
Qy 361 CAAGGTGCGCAAGAGAAACCAAGTAAGAGAACTCTCCCACTGACCCGGGACAT 420
Db 36239 CAAGATCAGGCAAGAGAAACCAATGTAAAGAGTAATCTCCCACTGACCCGGGTACAT 36298
Qy 421 AGCACCTCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCTCTCGTACC 480
Db 36299 GGCACCTCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCTCTCGTACC 36358
Qy 481 CATACTCGCTGTTGAGCCTATTTAATACCACTCTACTCGGCTCCATGAGGTCTCAGCC 540
Db 36359 CATACTCGCTGTTGAGCCTATTTAATACCACTCTACTCGGCTCCATGAGGTCTCAGCC 36418
Qy 541 CAAACCCCTACTGTTGATGTGCTCCCTGCACTTTCAGGCCATACATTTCAATC 600
Db 36419 CAAACCCCTACTGTTGATGTGCTCCCTGCACTTTCAGGCCATATGTTTCAATC 36478
Qy 601 CCTGTTCTGAACTAAGTGAACAACTTTCAGCACAGAAATAAACCACCTTCCGTTTTAGTA 660
Db 36479 CCTGTTCTGAACTAAGTGAACAACTTTCAGCACAGAAATAAACCACCTTCCGTTTTAGTA 36538
Qy 661 GGACCTCTGTTTCCAATCTGGAATAACCATACCTCAAACTCTCAAACTCTGTGTAAATTT 720
Db 36539 GGACCTCTGTTTCCAATCTGGAATAACCATACCTCAAACTCTCAAACTCTGTGTAAATTT 36598
Qy 721 AGCAATCTATAGACACACACAGCTCCCAATGCATCAGGTGGGTAAACCTCCCAACAGA 780
Db 36599 AGCAATCTATAGACACACACAGCTCCCAATGCATCAGGTGGGTAACTCTCTCCCAACAA 36658
Qy 781 ATAGTCTGCTACCTCAGGAATAATTTTGTGTGTGTACTCTCAGCTATCATTTGTTTG 840
Db 36659 ATAGTCTGCTACCTCAGGAATAATTTTGTGTGTGTACTCTCAGCTATCATTTGTTTG 36718
Qy 841 AATGGCTCTCAGAACTATATGTTCTCTCATTTCTTAGTGCCCTCTATGACCATCTAC 900
Db 36719 AATGGCTCTCAGAACTATATGTTCTCTCATTTCTTAGTGCCCTCTATGACCATCTAC 36778
Qy 901 ACTGAACAGATTTTATCAATCATGTCTGCTAGCCCAACCAAGAGTACCCATT 960
Db 36779 ACTGAACAGATTTTATACAGTTATGTCATATCTAAGCCCGGCAACAAAGAGTACCCATT 36838
Qy 961 CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGCAGACTAGGTACTTGGCATTTGGCAGTATC 1020
Db 36839 CTTCTCTTTTGTATCAGAGCAGGAGTGTAGGTGCACTAGGTACTTGGCATTTGGCAGTATC 36898
Qy 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAATATGGTGACATGACAG 1080
Db 36899 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAACTATATATGGGACATGGAACGG 36958
Qy 1081 GTCACTGACTCCCTCGTCCACTTGAAGATCAACTTAATCTCCCTAGCAGCAGTAGTCTT 1140

Db 36959 GTCCGCGACTCCCTGGTCACTTGTCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCCTT 37018
Qy 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTTAA 1200
Db 37019 CAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTAA 37078
Qy 1201 GGAGAAGAACGCTGTTTATTTATGTTAATCAATCCAGAAATGTCTCACTGAGAAAGTTAAAGAA 1260
Db 37079 GGGGAAGATGCTGTTTATTTATGTTAATCAATCCAGAAATGTCTCACTGAGAAAGTTAAAGAA 37138
Qy 1261 ATTCGAGATCGAATCAATGTAGAGCAGAGGAGCTTCAAAACCGAAGCTGGGCGCTC 1320
Db 37139 ATTCGAGATCGAATCAACAGTAGAGCAGAGGAGCTTCGAAACACTGGACCCCTGGGCGCTC 37198
Qy 1321 CTGAGCCATGGAATGCCCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAATATTG 1380
Db 37199 CTCAGCAATGGAATGCCCTGGTCTCCCTCTTAGGACCTTAGCAGCTCTAATATTG 37258
Qy 1381 TTACTCTCTTTTGGACCTCTATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGAAAT 1440
Db 37259 CTACTCTCTTTTGGACCTCTATCTTTAACTCTCTTGTAAAGTTTGTCTCTCCAGAAAT 37318
Qy 1441 GAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
Db 37319 GAAGCTGTAAACTACAAATGGAGCCCAAGATGCAGTCCAA 37359

RESULT 11
US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US2004004298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauclia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = a or g or c or t/u
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-30

Query Match 89.4%; Score 1324.2; DB 17; Length 7582;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 1365; Conservative 24; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATACTTTTCTTTTACTGTTCTTACCCCTTTTGGCTCTCACT 60
Db 5581 ATGGSCCTCCCTTATCATATTTTCTCTKTASGTTSTTTTACCTTTTCACTCTCACT 5640

Qy 61 GCACCCCTCCATGCTGCTGTACACACAGTAGTCCCTTACCAAGAGTTTCTATGAGA 120
Db 5641 GCACCCCTCCATGCTGCTGTATGACACAGTAGTCCCTTACCMAGAGTTTCTATGAGA 5700

Qy 121 ACGGGCTTCTCGGAATATTGATGCCCATATATAGGAGTTTATCTAGGGAACCTCC 180
Db 5701 ATGCAGCGTCCCGGAATATTGATGCCCATATGATAGGAGTTTCTTSTAAGGGAACCTCC 5760

Qy 181 ACCTTCACTGCCACACACCCATATGCTCCGCAACTGCTATAACTCTGCACTTTTGCATG 240
Db 5761 ACCTTCACTGCCACACCCATATGCTCCGCAACTGCTATACTCTGCACTTTTGCATG 5820

Qy 241 CATGCAAACTACTATTATTGGACAGGGAATTAATCTAGTTGTTCTGAGGACTT 300
Db 5821 CATGCAAACTACTATTATTGGACAGGGAATTAATCTAGTTGTTCTGAGGACTT 5880

Qy 301 GGAGCCACTGCTGTTGGACTTACCCATACAGTATGCTGATGGGGTGGAAAT 360
Db 5881 GGAGTCACTGCTGTTGGACTTACCCATACAGTATGCTGATGGGGTGGAAAT 5940

Qy 361 CAAGGTCAGGCAAGAGAAACAAAGTAAAGAGCAATCTCCCAACTGACCCCGGGGACAT 420
Db 5941 CAAGATCAGGCAAGAGAAACAAAGTAAAGAGCAATCTCCCAACTGACCCCGGGTACAT 6000

Qy 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCGTACC 480
Db 6001 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCGTACC 6060

Qy 481 CATACTCGCTGCTGAGCCCTATTATACACACCTCACTCGGCTCCATGAGGTCTCAGCC 540
Db 6061 CATACTCGCTGCTGAGCCCTATTATATACACACCTCACTCGGCTCCATGAGGTCTCAGCC 6120

Qy 541 CAAAACCTTACTAACTGTTGGATGCTCCCTCCCTGCACTTTCAGGCAATACATTTCAATC 600
Db 6121 CAAAACCTTACTAACTGTTGGATGCTCCCTCCCTGCACTTTCAGGCAATATGTTTCAATC 6180

Qy 601 CCGTTCCTGAAACAAATGGAACAACTTCAGCAGAGAAATAACACCACTTCGTTTATGTA 660
Db 6181 CCGTACCTGAAACAAATGGAACAACTTCAGCAGAGAAATAACACCACTTCGTTTATGTA 6240

Qy 661 GGACCTCTTGTTCCTCAATCTGGAATAACCCATACCTCAAACTCACTCTGTGTAATTT 720
Db 6241 GGACCTCTTGTTCCTCAATSTGGAATAACCCATACCTCAAACTCACTCTGTGTAATTT 6300

QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGACAG 1080
 DB 1783 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGACAG 1842
 QY 1081 GTCACTGACCTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1140
 DB 1843 GTGCCGACCTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1902
 QY 1141 CAANAATCGAAGAGCTTTAGACTTGTCAACCGCAAGAGGGGGAACCTGTTTATTTT 1200
 DB 1903 CGAANAATCGAAGAGCTTTAGACTTGTCAACCGCTGAGAGGGGGAACCTGTTTATTT 1962
 QY 1201 GGAAGAAGACGCTGTTTATTTATTAATCAATCAGAAATTCCTCACTGAGAAAGTTAAGAA 1260
 DB 1963 GGGGAAGAGCTGTTTATTTATTAATCAATCCGAATCTCACTGAGAAAGTTGAGAA 2022
 QY 1261 ATTCCAGATCGAATACAAATAGAGCAGAGAGCTTCAAAACACCGGAACCTGGGGCTC 1320
 DB 2023 ATTCCAGATCGAATACAAATAGAGCAGAGAGCTTCAAAACACCTGGAGCCTGGGGCTC 2082
 QY 1321 CTGAGCAATGGATGGCTGGTCTCCCTCTTAGGAGCTCTAGCAGCTCTAAATATTG 1380
 DB 2083 CTCAGCGATGGATGGCTGGTCTCCCTCTTAGGAGCTCTAGCAGCTCTAAATATTG 2142
 QY 1381 TTACTCTCTTTGGACCTGTATCTTTAACTCTCTTGAAGTTGTCTCTCCAGAAAT 1440
 DB 2143 CTACTCTCTTTGGACCTGTATCTTTAACTCTCTTGAAGTTGTCTCTCCAGAAAT 2202
 QY 1441 GAAGCTGTAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
 DB 2203 GAAGCTGTAAGCTACAAATGGAGCCCAAGATGCAATGCCAA 2243
 RESULT 13
 US-10-632-793-26
 ; Sequence 26, Application US/10632793
 ; Publication No. US20040048298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PARANHOS-BACCALA, Glauca
 ; APPLICANT: MALLET, Francois
 ; APPLICANT: VOISSET, Cecile
 ; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
 ; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
 ; FILE REFERENCE: 110048
 ; CURRENT APPLICATION NUMBER: US/10/632,793
 ; CURRENT FILING DATE: 2003-08-04
 ; PRIOR APPLICATION NUMBER: US/09/869,927
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: PCT/FR00/00144
 ; PRIOR FILING DATE: 2000-01-21
 ; PRIOR APPLICATION NUMBER: FR 99/00888
 ; PRIOR FILING DATE: 1999-01-21
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 26
 ; LENGTH: 2782
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-632-793-26
 Query Match 88.2%; Score 1306.6; DB 17; Length 2782;
 Best Local Similarity 92.6%; Pred. No. 0;
 Matches 1372; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
 QY 1 ATGGCCCTCCCTTATACATCTTTCTTTACTGTCTTACCCCTTTCCGCTCTCACT 60
 DB 763 ATGGCCCTCCCTTATCATATTTCTCTGTAGTGTCTTCCACCTGTTTCACTCTCACT 822
 QY 61 GCACCCCTCCATGCTGCTGACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
 DB 823 GCACCCCTCCATGCTGCTGATGACCAAGTAGTCCCTTACCAAGAGTTTCTATGGAGA 882
 QY 121 ACGCGGCTTCTGGAATATTGATGCCCCCATATAGGAGTTTATCTAAGGGAAACTCC 180

DB 883 ATGACGCGTCCGGAATAATTGATGCCCATCGTAGGAGTCTTTCTAAGGGAACCCCC 942
 QY 181 ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATAAATCTCTCCACTCTTTGCGATG 240
 DB 943 ACCTTCACCTGCCACACCCCATATGCCCGCAACTGCTATCACTCTGCGCACTCTTTGCGATG 1002
 QY 241 CATGCAAAATCACTCATTTATTGGACAGGAAATGATTAACTCTAGTTGTCTGCGAGGACTT 300
 DB 1003 CATGCAAAATCACTCATTTATTGGACAGGAAATGATTAACTCTAGTTGTCTGCGAGGACTT 1062
 QY 301 GAGGCCACTGTCTGTTGGACTTACTTACCCATACAGTATGCTCTGATGGGGTGGAAAT 360
 DB 1063 GAGGTCACTGTCTGTTGGACTTACTTACCCAACTGGTATGCTGATGGGGTGGAGTT 1122
 QY 361 CRAAGTCAGGCAAGAGCAAAAAAAGTAAAGAGCAATCTCCCAACTGACCCCGGGGACAT 420
 DB 1123 CRAAGTCAGGCAAGAGCAAAAAAAGTAAAGAGTAAATCTCCCAACTGACCCGGGTACAT 1182
 QY 421 AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAAACTACATGAAACCTCCGTTACC 480
 DB 1183 GGCACCTCTAGCCCCCTACAAAGGACTAGATCTCTCAAAACTACATGAAACCTCCGTTACC 1242
 QY 481 CATACTCGCTGGTGGAGCTTATTAATACACCTCACTCGGCTCCATGAGGTCTCAGCC 540
 DB 1243 CATACTCGCTGGTGGAGCTTATTAATACACCTCACTCGGCTCCATGAGGTCTCAGCC 1302
 QY 541 CAAAACCTCACTAACTGTTGGATGCTCCCTCCCTGACCTTCAGGGCCATACATTTCAATC 600
 DB 1303 CAAAACCTCACTAACTGTTGGATGCTCCCTCCCTGACCTTCAGGGCCATGATGTTCAATC 1362
 QY 601 CTTGTTCTGTAACAATGGAACCAACTTCAGCACAGAAATAAACCACTTCCTGTTTATGTA 660
 DB 1363 CTTGTTCTGTAACAATGGAACCAACTTCAGCACAGAAATAAACCACTTCCTGTTTATGTA 1422
 QY 661 GGAACCTCTGTTTCCAATCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATAATTT 720
 DB 1423 GGAACCTCTGTTTCCAATCTGGAATAAACCCATACCTCAAACTCAGCTGTGTAATAATTT 1482
 QY 721 AGCAATACATACACACACACCTCCCAATGATCAGTGGGTGTAACACCTCCACACGA 780
 DB 1483 AGCAATACATACACACACACCTCCCAATGATCAGTGGGTGTAACCTCCACACGA 1542
 QY 781 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTTGTTACCTCAGCTATCATTTGTTG 840
 DB 1543 ATAGTCTGCTACCTCAGGAATAATTTTGTCTGTTGTTACCTCAGCTATCATTTGTTG 1602
 QY 841 AATGGCTCTTCAGAAATCTATGCTCTCTCAATCTTATGTCCTCCCTATGACCATCTAC 900
 DB 1603 AATGGCTCTTCAGAAATCTATGCTCTCTCAATCTTATGTCCTCCCTATGACCATCTAC 1662
 QY 901 ACTGAAACAAGATTTATACATCATGCTGTAACCTTAAGCCCAACAAAGAGTACCCATT 960
 DB 1663 ACTGAAACAAGATTTATACATCATGCTGTAACCTTAAGCCCAACAAAGAGTACCCATT 1722
 QY 961 CTTCTCTTTGTTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGTCGATTTGGCAGTATC 1020
 DB 1723 CTTCTCTTTGTTATAGGAGCAGGAGTCTAGGTGCACTAGGTACTGCACTTTGGCGGTATC 1782
 QY 1021 ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGACAG 1080
 DB 1783 ACAACCTCTACTCAGTTTCTACTACAACTATCTCAAGAAATAAATGGTGACATGGACAG 1842
 QY 1081 GTCACTGACCTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1140
 DB 1843 GTGCCGACCTCCCTGGTCACTTGTGAAGATCAACTTAACTCCCTAGCAGAGTAGTCTT 1902
 QY 1141 CAANAATCGAAGAGCTTTAGACTTGTCAACCGCAAGAGGGGGAACCTGTTTATTTT 1200
 DB 1903 CGAANAATCGAAGAGCTTTAGACTTGTCAACCGCTGAGAGGGGGAACCTGTTTATTT 1962
 QY 1201 GGAAGAAGACGCTGTTTATTTATTAATCAATCAGAAATTTGCTACTGAGAAAGTTAAGAA 1260

Db 1963 GGGGAAGATGCTGTTATTTATGTTAACTCAATCCGGAATCGTCACTGAGAAAGTTGAAGA 2022

Qy 1261 ATTGAGATCGAATACAATGTAGACAGAGGAGCTTTCAAAACACCGAAGCTGGGGCTTC 1320

Db 2023 ATTCCAGATCGAATACAACGTATAGCAGAGGAGCTTCGAAACACATGGACCCCTGGGGCTTC 2082

Qy 1321 CTCAGCAATGATGCCCTGGGTTCTCCCTTCTTTAGGACCTCTAGCAGCTCTTAATATTG 1380

Db 2083 CTCAGCCGATGATGCCCTGGGATTCTCCCTTCTTTAGGACCTCTAGCAGCTCTTAATATTG 2142

Qy 1381 TTACTCCTCTTTGGACCCCTGTATCTTTAACTCCCTTGTAAAGTGTGTCTCTTCCAGAAAT 1440

Db 2143 CTACTCCTCTTTGGACCCCTGTATCTTTGACCTCCTTGTAACTTTGTCTCTTCCAGAAATC 2202

Qy 1441 GAAGCTGTAAAGCTACAGATGTCTTTACAAATGGAAACCCCA 1481

Db 2203 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCCAA 2243

RESULT 14

US-10-087-192-730/c

; Sequence 730, Application US/10087192

; Publication No. US20020182586A1

; GENERAL INFORMATION:

; APPLICANT: Morris, David W.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

; TITLE OF INVENTION: CANCER

; FILE REFERENCE: 529452000122

; CURRENT APPLICATION NUMBER: US/10/087,192

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 2059

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 730

; LENGTH: 161334

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)...(161334)

; OTHER INFORMATION: n = A,T,C or G

US-10-087-192-730

Query Match 86.3%; Score 1277.4; DB 13; Length 161334;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 1399; Conservative 0; Mismatches 81; Indels 33; Gaps 4;

Qy 1 ATGGCCCTCCCTATCATATCTTTTCTCTTTACTGTTCTCTTACCCCTTTGCTCTCACT 60

Db 102037 ATGGCCCTCCCTATCATATCTTTTCTCTTTACTGTTCTCTTACCCCTTTTCACTCTCACT 101978

Qy 61 GCACCCCTCCATGCTGTACAACAGTAGCTCCCTTACCAGAGTTCTCTATGAAGA 120

Db 101977 GCACCCCTCCATGCTGTACAACAGTAGCTCCCTTACCAGAGTTCTCTATGAAGA 101918

Qy 121 ACGCGGCTTCTCGAAATATTGATGCCCCCATCATATAGGAGTTTATCTAAGGGAACTCC 180

Db 101917 ATGCGGCTTCCCAAGAAATATTGATGCCCATCAATAGGAGTTTACCTAAGGAACTCC 101858

Qy 181 ACCTTCAGTCCCAACCCATATGCCCGCACTGCTATACTCCCACTCTTTGCATG 240

Db 101857 ACCTTCAGTCCCAACCCATATGCCCGCACTGCTATACTCTCCCACTCTTTGCATG 101798

Qy 241 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGTCTGAGGACTT 300

Db 101797 CATGCAATACTCATTTATGGACAGGAAATGATTAATCTAGTTGTCTGAGGACTT 101738

Qy 301 GGAGCCACTGTCTGTGGACTTACTTCAACCATACAGATATGTCTGATGGGGTGAATTT 360

Db 101737 GGAGCCACTGTCTGTGCGACTTACTTCAACCCACTATGTTATGTCTGAGGGGGTGGAGTT 101678

Qy 361 CAAGGTGAGCAAGAGAAACAAGTAAGGAAGCAATCTCCCACTGACCCGGGACAT 420

Db 101677 CAAGATCAGGCAAGAGAAACAAGTAAGGAAGCAATCTCCCACTGACCCGGGTAGAT 101618

Qy 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTC-AAAACTACATGAAACCCCTCCGTAC 479

Db 101617 AGCACCCCTAGCCCTTACAAAGGACTAGATCTCTCTAAACTACATGAAACCCCTCCATAC 101558

Qy 480 CCATACCTGCGCTGTGAGCCCTATTAAATAACAACCTCTCGGTCTCATAGGTCTCAGC 539

Db 101557 CCATACCTGCGCTGTGAGCCCTATTAAATACCACTCTCAGGTCTCCTGAGGTCTCGGC 101498

Qy 540 CCAAAACCCCTACTAATCTTGGATGTGCTCCCTGCTGCTTCAAGCCATACATTTCAAT 599

Db 101497 CCAAAACCCCTACTAATCTTGGATGTGCTCCCTGCTGCTTCAAGCCATGCAATTTCAAT 101438

Qy 600 CCCTGTTCTCGAAACAATGGAAACAATCTCAGCACAGAAATAAACACCACTTCGGTTTTAGT 659

Db 101437 CCCTGTACTCTGAAACAATGGAAACAATCTCAGCACAGAAATAAACACCACTTCGGTTTTAGT 101378

Qy 660 AGGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAAACTCTCACTGTGTAAAAAT 719

Db 101377 AGGACCTCTTGTTCCTCAATCTGGAAATAACCCATACCTCAAACTCTCACTGTGTAAAAAT 101318

Qy 720 TAGCAATACTATAGACACAACAGCTCCCAATGCAATGAGTGGGTAAACACCTCCACACAG 779

Db 101317 TAGCAATACTATAGACACAACAGCTCCCAATGCAATGAGTGGGTAACTCTCCACACAG 101258

Qy 780 AATAGTCTGCTTACCCCTCAGGAATATTTTTGTCTGTGTACCTCAGCCTATCATTTGTTT 839

Db 101257 AATAGTCTGCTTACCCCTCAGGAATATTTTTGTCTGTGTACCTTACCTATCGTTGTTT 101198

Qy 840 GAATGGCTCTTCAGAAATCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 899

Db 101197 GAATGGCTCTTCAGAAATCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 101139

Qy 900 CACTGAACAAGATTTTATACAATCATGTCGTACCTAAGCCCAACAAACAAAGAGTACCCAT 959

Db 101138 CACTGAACAAGATTTTATACAATCATGTCGTACCTAAGCCCAACAAACAAAGAGTACTCAT 101079

Qy 960 TCTTCTCTTTTGTATCAGACGAGGAGTCTAGGACAGCTAGGTACTGCGCTGGCAGTAT 1019

Db 101078 TCTTCTCTTTTGTATCAGACGAGGAGTCTAGGTGCTAGGTCTGCGCTGGCAGTAT 101019

Qy 1020 CACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAATGTGACATGGAACA 1079

Db 101018 CACAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAATAATGTGACATGGAATG 100959

Qy 1080 GGTCACTGACTCCCTGGTCACTTTCGAAAGATCAACTTAACTCCCTAGCAGCAGTACTCT 1139

Db 100958 GGTTCGCACTCCCTGGTCACTTTCGAAAGATCAACTTAACTCCCTAGCAGCAGTACTCT 100899

Qy 1140 TCATAATCGAAGAGCTTTAGACTTCTCTAAACCGCCAAAGAGGGGAACTCTGTTATTTT 1199

Db 100898 TCATAATCGAAGAGCTTTAGACTTCTCTAAACCTCTGAAAGAGGGGAACTCTGTTATTTT 100839

Qy 1200 AGGAGAAAGACGCTG-----TTATATGTTAAATCA 1229

Db 100838 AGGAGAAAGATGTTTATTTATTTAGCGAAGAAATGTTGTTATTTATGTTAAATCA 100779

Qy 1230 ATCCAGAAATGTGCTAGAGAAAGTTTAAAGAAATTCGAGATCGAATCAATGTAGAGCAGA 1289

Db 100778 ATCCGAAATGTGCTAGAGAAAGTTTAAAGAAATTCGAGATCGAATCAATGTAGAGCAGA 100719

Qy 1290 GGAGCTTC-AAAAACCGAAGCCTGGGCTCTCTCAGCCAAATGGATGCCCTGGGTTCTCC 1348

Db 100718 GGAGCTTC-AAAAACCGAAGCCTGGGCTCTCTCAGCCAAATGGATGCCCTGGGTTCTCC 100659

Qy 1349 CCTTCTTAGGACCTCTAGGAGCTCTAATATTTGTTACTCTCTCTTGGACCCCTGATCTTTA 1408

Db 100658 CCTTCTTAGGATCTCTAGGAGCTCTAATATTTGTTACTCTCTCTTGGACCCCTGATCTTTA 100599

D _b	885	ATGCTGACATGGAAATGGGTCCGCTGATAACCGTGTGCACCTTGTCGAAGATCAACTTAATCCTCC	944
Q _y	1124	TAGCAGCAGTAGTCCTTTCAAATCGAAGAGCTTTAGACTTGTCTAAACGCCCAAAGAGGGG	1183
D _b	945	TAGCAGCAGTAGTCCTTTCAAANATCGAAGAGCTTTAGACTTGTCTAAACGCCGAAGAGCGGG	1004
Q _y	1184	GAACCTCTTTATTATTTTAGGAGAAGACGCTGTTATTATGTTAAATCAATTCAGAAATTTGTCA	1243
D _b	1005	GAACCTTTTATTTTAGAGGAAAATGCTGTGTGTTATGTTAAATCAATCCGGAATCATCA	1064
Q _y	1244	CTGAGAAAGTTAAAGAAATTCGAGATCGAATCAAACTAGCAGAGAGGAGCTTCAAACA	1303
D _b	1065	CCGAGAAAGTTAAAGAAATTCAGGTGCGAATAAATACGTAGCAGAAAGGAGCTGCCAAACA	1124
Q _y	1304	CCGAACGCTGGGGCTCCTCAGCCAAATGGATGCCCTGGGTTCTCCCCTCTTTAGGACCTC	1363
D _b	1125	CTGACCCTGGGGCTCCTTAGCCAAATGGATGCCCTGGATTCTCCCCCTCTTAGGACCTC	1184
Q _y	1364	TAGCAGCTCTAAATATGTTACTCTCTTTGGAACCTGTACTTCTTAAACCTCCTCTGTTAAGT	1423
D _b	1185	TAGCAGCTATAATATGTTACTCTCTTTGGACCTGTACTCTTTAAACCTCCTCTGTTAAGT	1244
Q _y	1424	TTGTCTCTTCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAAACCCCA	1481
D _b	1245	TTGTCTTTTCAGAAATCGAAGCAGTAAACTTCAAAATCGTCTTTCAAAATGGAGCCCCA	1302

RESULT 16

```

US-10-632-793-24
; Sequence 24, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (84)..(84)
; OTHER INFORMATION: n = a o r g o r c o r t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (193)..(193)
; OTHER INFORMATION: n = a o r g o r c o r t/u
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (241)..(241)
; OTHER INFORMATION: n = a o r g o r c o r t/u
; US-10-632-793-24

```

QY	61	GCACCCCTCCATGCTGCTGTACACACAGTAGTCTCCCTTACCAAGAGTGTCTCTATGAAGA	120
Db	779	GCACCCCTCCATGCGGCTGTATGACCAAGTAGTCTCCCTTACCAAGAGTGTCTCTATGAAGA	838
QY	121	ACGCGCTTCTCGGAAATATTGATGCCCCATCATATAGGAGTTTTATCTAAGGGAACCTCC	180
Db	839	ATGCAGCTCCCGGAATATTGATGCCCCATCGTAGAGTCTTTGTGAAGGAACCCCC	898
QY	181	ACCTTCAGTGCCCAACCCATATGCCCGGAACGTCTATACTCTGCCACCTCTTTGCATG	240
Db	899	ACCTTCAGTGCCCAACCCATATGCCCGGAACGTCTATACTCTGCCACCTCTTTGCATG	958
QY	241	CATGCAAACTACTATTATGGACAGGGAATAATGATTAATCTAGTTGTCTTGGAGGACTT	300
Db	959	CATGCAAACTACTATTATGGACAGGGAATAATGATTAATCTAGTTGTCTTGGAGGACTT	1018
QY	301	GGAGCCACTGTCTGTGGACTTACTTACCCTCATACAGTATGTCTGATGGGGTGGAAATT	360
Db	1019	GGAGTCACTGTCTGTGGACTTACTTACCCTCATACAGTATGTCTGATGGGGTGGAGATT	1078
QY	361	CAAGGTCAGCAGAGAGAAAAACAAGTAAGGAAGCAATCTCCAACTCACCCGGGGACAT	420
Db	1079	CAAGATCAGCAAGAGAAAAACATGTAAAGAAGTAAATCTCCCACTCACCCGGGTACAT	1138
QY	421	AGCACCCCTAGCCCCCTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCGTACC	480
Db	1139	GGACCTCTAGCCCCCTACAAAGGACTAGATTCTCTCAAACTACATGAACCCCTCGTACC	1198
QY	481	CATAGTCGCTGTGAGCCCTTTAATACACCCCTCACTCGCTCCATGAGGTCTCACGCC	540
Db	1199	CATAGTCGCTGTGAGCCCTTTAATACACCCCTCACTCGCTCCATGAGGTCTCGGCC	1258
QY	541	CAAAACCTACTAACTGTTGGATGTCCTCCCTCGCTCAGGCCCATACATTTCAATC	600
Db	1259	CAAAACCTACTAACTGTTGGATGTCCTCCCTCGCTCAGGCCCATATGTTTCAATC	1318
QY	601	CCTGTTCTGAAACAATGGAAACAATTCAGCAGAGAAATAAACAACACTTCCTTTTAGTA	660
Db	1319	CCTGTAACCTGAAACAATGGAAACAATTCAGCAGAGAAATAAACAACACTTCCTTTTAGTA	1378
QY	661	GGACCTCTGTTTCCCAATCTGGAATAATAACCATCTCAAACTCTCACCTGTGTAAATTT	720
Db	1379	GGACCTCTGTTTCCCAATCTGGAATAATAACCATCTCAAACTCTCACCTGTGTAAATTT	1438
QY	721	AGCAATACTATAGACACAACACAGCTCCCAATGCATCAGTGGGTAAACCTCCACACGA	780
Db	1439	AGCAATACTATAGACACAACACAACTCCCAATGCATCAGTGGGTAACTCTCCACACAA	1498
QY	781	ATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCCTATCATTTGTTG	840
Db	1499	ATAGTCTGCCCTACCTCAGGAATATTTTGTCTGTGGTACCTCAGCCCTATCGTTGTTG	1558
QY	841	AATGGCTCTTCAGAACTATATGTGCTTCTCTCAATCTTTAGTGCCCTCATGACCACTAC	900
Db	1559	AATGGCTCTTCAGAACTATATGTGCTTCTCTCAATCTTTAGTGCCCTCATGACCACTAC	1618
QY	901	ACTGAACAAGATTTTATACATATGTGCTACCTAAGCCCAACAACAAAGATGACCAATT	960
Db	1619	ACTGAACAAGATTTTATACAGTTATGTATATCTTAAGCCCGCAACAACAAAGATGACCAATT	1678
QY	961	CTTCTCTTTGTTATCAGACGAGGAGTGCTAGGCAGACTAGGTACTGGCATTTGGCAGTATC	1020
Db	1679	CTTCTCTTTGTTATAGGACGAGGAGTGCTAGGTGCATAGGTACTGGCATTTGGCGTATC	1738
QY	1021	ACAACTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGGTGAACGACAG	1080
Db	1739	ACAACTCTACTCAGTTTCTACTACAAACTATCTCAAGAAATAAATGGGACATGGAACGG	1798
QY	1081	GTCACTGACTCCCTGGTCACTCTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT	1140
Db	1799	GTGCGCGGACTCCCTGGTCACTCTGCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTT	1858

QY 1141 CAAATCGAAGAGCTTTAGACTTGTCTAAACGCGCAAGAGGGGAAACCTGTTTATTTTAA 1200
Db |||||||
QY 1859 CAAATCGAAGAGCTTTAGACTTGTCTAAACGCGTGAAGAGGGGGAACCTGTTTATTTTAA 1918
Db |||||||
QY 1201 GGAGAGAAGAGCTGTTTATTATGTTAAATCAA 1230
Db |||||||
QY 1919 GGGGAAGAAGCTGTTATTATGTTAAATCAA 1948
Db |||||||

RESULT 17
US-09-864-761-21192
; Sequence 21192, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIORITY APPLICATION NUMBER: US 60/180,312
; PRIORITY FILING DATE: 2000-02-04
; PRIORITY APPLICATION NUMBER: US 60/207,456
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY APPLICATION NUMBER: US 09/632,366
; PRIORITY FILING DATE: 2000-08-03
; PRIORITY APPLICATION NUMBER: GB 24263.6
; PRIORITY FILING DATE: 2000-10-04
; PRIORITY APPLICATION NUMBER: US 60/236,359
; PRIORITY FILING DATE: 2000-09-27
; PRIORITY APPLICATION NUMBER: PCT/US01/00666
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00667
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00664
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00669
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00665
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00668
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00663
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00662
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00661
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: PCT/US01/00670
; PRIORITY FILING DATE: 2001-01-30
; PRIORITY APPLICATION NUMBER: US 60/234,687
; PRIORITY FILING DATE: 2000-09-21
; PRIORITY APPLICATION NUMBER: US 09/608,408
; PRIORITY FILING DATE: 2000-06-30
; PRIORITY APPLICATION NUMBER: US 09/774,203
; PRIORITY FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21192
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
; OTHER INFORMATION: NT HIT: AF208161.1, EVALUATE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: AU138405.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUATE 2.00e-16
US-09-864-761-21192

Query Match 48.4%; Score 716.6; DB 9; Length 792;
Best Local Similarity 94.7%; Pred. No. 1.4e-209;
Matches 753; Conservative 0; Mismatches 39; Indels 3; Gaps 1;
QY 375 AGAAAAACAAGTAAGGAAGCAATCTCCCAACTGACCCCGGGGACATAGACACCCCTAGCCCC 434
Db |||||||
QY 435 CTACAAAGAGCTAGTCTCTCAAAAGTACATGAAACCCCTCCGTAACCATCTCGCTCGCTGGT 494
Db |||||||
QY 61 CTACAAAGAGCTAGATCTCTCAAAAGTACATGAAACCCCTCCATACCCATCTAGCTGGCTGGT 120
Db |||||||
QY 495 GAGCCTATTATATACCAACCTCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCTACTATA 554
Db |||||||
QY 121 AAGCCTATTATATACCAACCTCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCTACTATA 180
Db |||||||
QY 555 CTGTTGGATGTGCTCCCTCCCTCGACCTTCAGGCCCATACATTTCAATCCCTGTTCTCTGAACA 614
Db |||||||
QY 181 CTGTTGGATGTGCTCCCTCCCTCGACCTTTAGGCCATACATTTCAATCCCTATACCTGAACA 240
Db |||||||
QY 615 ATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGTTTGTAGTAGGAGCTCTTGTTC 674
Db |||||||
QY 241 ATGGAACAACCTTCAGCACAGAAATAAACACCACTTCGTTTGTAGTAGGAGCTCTTGTTC 297
Db |||||||
QY 675 CAATCTGGAATAAACCACTACCTCAAAACCTCACTGTGTAAATTTTAGCAATACTATAGA 734
Db |||||||
QY 298 CAATCTGGAATAAACCACTACCTCAAAACCTCACTGTGTAAATTTTAGCAATACTATAGA 357
Db |||||||
QY 735 CACAACAGCTCCCAATGCAATCAGGTGGGTAAACACCTCCCAACGAATAGTGTGCTTACC 794
Db |||||||
QY 358 CACAGCCAACTCCCAATGCAATCAGGTGGGTAACTCCTCCCAACGAATAGTGTGCTTACC 417
Db |||||||
QY 795 CTGAGGAATATTTTGTCTGTGTGCTCAGCTTATCATTTGTTGATGCTCTTCA 854
Db |||||||
QY 418 CTGAGGAATATTTTGTCTGTGTGCTCAGCTTATCATTTGTTGATGCTCTTCA 477
Db |||||||
QY 855 ATCTATGTGCTCTCTCTCATTTCTAGTGCCCTTATGACCATCTACACTGAAACAAGATTT 914
Db |||||||
QY 478 ATCTGTGTGCTCTCTCTCATTTCTAGTGCCCTTATGACCATCTACACTGAAACAAGATTT 537
Db |||||||
QY 915 ATCAATCATGTGCTTACCTAAGCCCAACAAAGAGTACCCATTTCTTTTGTAT 974
Db |||||||
QY 538 ATCAATCATGTGCTTACCTAAGCCCAACAAAGAGTACCCATTTCTTTTGTAT 597
Db |||||||
QY 975 CAGAGCAGGAGTGTCTAGGAGAGCTAGGTACTGGCATTTGGCAGTATCAACCTCTACTCA 1034
Db |||||||
QY 598 TGGAGCAGGAGTGTCTAGGCGGAGTAGCTACTGCGATTTGGCGGTATCAACCTCTACTCA 657
Db |||||||
QY 1035 GTTCTACTCAAACTATCTCAAGAAATAAATGGTGAATGGAACAGGTCACTGACTCCCT 1094
Db |||||||
QY 658 GTTCTACTCAAACTGTCTCAAGAAATAAATGGTGAATGGAATGGGTGCGTGATACCT 717
Db |||||||
QY 1095 GGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGC 1154
Db |||||||
QY 718 GGTCACTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGC 777
Db |||||||
QY 1155 TTTAGACTTGTCTAAC 1169
Db |||||||
QY 778 TTTAGACTTGTCTAAC 792
Db |||||||

RESULT 18
US-10-363-616-228
; Sequence 228, Application US/10363616

Fr1 Feb 25 16:26:32 2005

```

; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; PRIOR FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 228
; LENGTH: 1684
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (89)...(1684)
; US-10-363-616-228

```

```

Query Match 40.0%; Score 591.8; DB 17; Length 1684;
Best Local Similarity 98.8%; Pred. No. 6.2e-171;
Matches 596; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 879 AGTCCCTCTGACCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCC 938
DB 721 AGTCCCTCTGACCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCC 780
QY 939 CCAACAACAAAGAGTACCCATCTCTCTTTTGTATCAGAGCAGGAGTCTAGGAGACT 998
DB 781 CCAACAACAAAGAGTACCCATCTCTCTTTTGTATCAGAGCAGGAGTCTAGGAGACT 840
QY 999 AGTACTGGCATTGGCAGTATCAACAACCTCTACTAGTTCTACTACAAATCTCTCAAGA 1058
DB 841 AGTACTGGCATTGGCAGTATCAACAACCTCTACTAGTTCTACTACAAATCTCTCAAGA 900
QY 1059 AATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTCCCAAGATCAACTTAA 1118
DB 901 AATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTCCCAAGATCAACTTAA 960
QY 1119 CTCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCAACGCCAAAAG 1178
DB 961 CTCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCAACGCCAAAAG 1020
QY 1179 AGGGGAACCTGTTATTTTAGAGAGAACCTGTTATATGTTAATCAATCCAGAAT 1238
DB 1021 AGGGGAACCTGTTATTTTAGAGAGAACCTGTTATATGTTAATCAATCCAGAAT 1080
QY 1239 TGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCA 1298
DB 1081 TGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCA 1140
QY 1299 AAACAACGAACGCTGGGCTCTCTAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGG 1358
DB 1141 AAACAACGAACGCTGGGCTCTCTAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGG 1200
QY 1359 ACCTTAGCAGCTCTAATATGTTACTCTCTTGGACCTGTATCTTTAACTCTCTTGT 1418
DB 1201 ACCTTAGCAGCTCTAATATGTTACTCTCTTGGACCTGTATCTTTAACTCTCTTGT 1260
QY 1419 TAAGTTTGTCTCTTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAAC 1478
DB 1261 TAAGTTTGTCTCTTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTCAAAATGGAAC 1320
QY 1479 CCA 1481
DB 1321 CCA 1323

```

```

RESULT 19
US-10-632-793-19
; Sequence 19, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia

```

```

; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 19
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-632-793-19

```

```

Query Match 39.4%; Score 582.8; DB 17; Length 591;
Best Local Similarity 99.7%; Pred. No. 2e-168;
Matches 584; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 887 CTATGACCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCCCAACA 946
DB 1 CCAATGGCCATCTACACTGAACAAGATTTATACAATCATGTGCTACCTAAGCCCAACA 60
QY 947 AAAGAGTACCCATCTCTCTTTTGTATCAGAGCAGGAGTCTAGGAGACTAGTACTG 1006
DB 61 AAAGAGTACCCATCTCTCTTTTGTATCAGAGCAGGAGTCTAGGAGACTAGTACTG 120
QY 1007 GCATGGGAGTATCAACAACCTCTACTCAGTTCTACTACAAATCTCTCAAGAAATAAATG 1066
DB 121 GCATGGGAGTATCAACAACCTCTACTCAGTTCTACTACAAATCTCTCAAGAAATAAATG 180
QY 1067 GTGACATGGAACAGGTCACTGACTCCCTGGTCACTTCCCAAGATCAACTTAACTCCCTAG 1126
DB 181 GTGACATGGAACAGGTCACTGACTCCCTGGTCACTTCCCAAGATCAACTTAACTCCCTAG 240
QY 1127 CAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGGAA 1186
DB 241 CAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGGAA 300
QY 1187 CCGTGTATTTTAGAGAGAAACGCTGTTATATGTTAATCAATCCAGAAATTTGCTACTG 1246
DB 301 CCGTGTATTTTAGAGAGAAACGCTGTTATATGTTAATCAATCCAGAAATTTGCTACTG 360
QY 1247 AGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCG 1306
DB 361 AGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGAGCTTCAAAACACCG 420
QY 1307 AACGCTGGGCTCTCTAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGGACCTCTAG 1366
DB 421 AACGCTGGGCTCTCTAGCCAAATGGATGCGCTGGGTTCTCCCTTCTTAGGACCTCTAG 480
QY 1367 CAGCTCTAATATTTTACTCTCTTTGGACCTCTGTATCTTTAACTCTCTTTGTTAAGTTTG 1426
DB 481 CAGCTCTAATATTTTACTCTCTTTGGACCTCTGTATCTTTAACTCTCTTTGTTAAGTTTG 540
QY 1427 TCTCTTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAAAT 1472
DB 541 TCTCTTCCAGAAATTTGAAGCTGTAAAGCTACAGATGGTCTTACAAAT 586

```

```

RESULT 20
US-10-632-793-25
; Sequence 25, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia

```



```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322491
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(570)
; OTHER INFORMATION: n = A,T,C or G
;
US-10-027-632-322492

Query Match      25.5%; Score 377.8; DB 13; Length 570;
Best Local Similarity 90.7%; Pred. No. 3e-105;
Matches 411; Conservative 2; Mismatches 39; Indels 1; Gaps 1;

QY      1  ATGGCCCTCCCTTATCATACATCTTTCTCTTTACTGTTCTCTTACCCCTTTGCTCTCACT 60
DB      463 ATAGCCCTCCCTTATCATATTTTCTTTTACCAITCTCTTACCTGCTTTTCACTCTCACT 404

QY      61  GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB      403 TCACCTCACTCCATGCTGTGTATGACACAGTAGTCTCCCTTACCAAGAGTTTCT-TGGAGA 345

QY      121 AC CGCGCTTCTCGGAATAATTTGATGCCCATATAGGAGTTTATCTAAGGGAATCTCC 180
DB      344 ATGCAGCTTCTCGGAATAATTTGATGCCCATATAGGAGTTTATCTAAGGGAATCTAC 285

QY      181 ACCTTCACCTGCCACACCCATATGCCCGCACTGTATAACTCTGCCACTCTTTTGCATG 240
DB      284 ACCTTCACCTGCCACACCCATATGCCCTTACCACTGTATAAGCTTCCACTCTTTTGCATG 225

QY      241 CATGCAAACTCACTATTATGGACAGGGAATAATGATTAATCTCTAGTTGCTCTGGAGACTT 300
DB      224 CATGCAAACTCACTATTATTTGTCATGGAAAAATGATTAATCTCTCTGGAGACTT 165

QY      301 GGAGCCACTGCTGTGTGGACTTACTTACCCTACCATACAGTAGTCTGTATGGGGGTGGAATT 360
DB      164 GGAGCCACTGCTGTGTGGACTTACTTACCCTACCATACAGTAGTCTGTATGGGGGTGGAATT 105

QY      361 CAAAGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGACAT 420
DB      104 CAAAGTCAGGCAAGAGAAAAACAAGTAAAGGAAGTAAATCTCCCAACTGACCCGGTACAT 45

QY      421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTC 453
DB      44  AGCACCCCTAGACCCCTACAAAGGACTAATCCCC 12

RESULT 24
US-10-027-632-322491/c
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322491
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(570)
; OTHER INFORMATION: n = A,T,C or G
;
US-10-027-632-322491

Query Match      25.5%; Score 377.8; DB 13; Length 570;
Best Local Similarity 90.7%; Pred. No. 3e-105;
Matches 411; Conservative 2; Mismatches 39; Indels 1; Gaps 1;

QY      1  ATGGCCCTCCCTTATCATACATCTTTCTCTTTACTGTTCTCTTACCCCTTTGCTCTCACT 60
DB      463 ATAGCCCTCCCTTATCATATTTTCTTTTACCAITCTCTTACCTGCTTTTCACTCTCACT 404

QY      61  GCACCCCTCCATGCTGTGTACAAACAGTAGTCTCCCTTACCAAGAGTTTCTATGAAGA 120
DB      403 TCACCTCACTCCATGCTGTGTATGACACAGTAGTCTCCCTTACCAAGAGTTTCT-TGGAGA 345

QY      121 AC CGCGCTTCTCGGAATAATTTGATGCCCATATAGGAGTTTATCTAAGGGAATCTCC 180
DB      344 ATGCAGCTTCTCGGAATAATTTGATGCCCATATAGGAGTTTATCTAAGGGAATCTAC 285

QY      181 ACCTTCACCTGCCACACCCATATGCCCGCACTGTATAACTCTGCCACTCTTTTGCATG 240
DB      284 ACCTTCACCTGCCACACCCATATGCCCTTACCACTGTATAAGCTTCCACTCTTTTGCATG 225

QY      241 CATGCAAACTCACTATTATGGACAGGGAATAATGATTAATCTCTAGTTGCTCTGGAGACTT 300
DB      224 CATGCAAACTCACTATTATTTGTCATGGAAAAATGATTAATCTCTCTGGAGACTT 165

QY      301 GGAGCCACTGCTGTGTGGACTTACTTACCCTACCATACAGTAGTCTGTATGGGGGTGGAATT 360
DB      164 GGAGCCACTGCTGTGTGGACTTACTTACCCTACCATACAGTAGTCTGTATGGGGGTGGAATT 105

QY      361 CAAAGTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGACCCGGGACAT 420
DB      104 CAAAGTCAGGCAAGAGAAAAACAAGTAAAGGAAGTAAATCTCCCAACTGACCCCTGGTACAT 45

QY      421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTC 453
DB      44  AGCACCCCTAGACCCCTACAAAGGACTAATCCCC 12

RESULT 23
US-10-027-632-322492/c
; Sequence 322492, Application US/10027632
```



```

; Sequence 322491, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 322491
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(570)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-322491

```

```

Query Match      25.5%; Score 377.8; DB 17; Length 570;
Best Local Similarity 90.7%; Pred. No. 3e-105;
Matches 411; Conservative 2; Mismatches 39; Indels 1; Gaps 1;

Qy 1 ATGGCCCTCCCTATCATCTTTCTCTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60
Db 463 ATAGCCCTCCCTATCATATTTTCTCTTACCATTCTCTTACCTGCTTTCACTCTCACT 404

Qy 61 GCACCCCTCCATGCTGCTACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 403 TCACCTACTCCATGCTGTTGATGACCAAGTAGCTCCCTTACCAAGAGTTTCT-TGGAGA 345

Qy 121 ACGCGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 180
Db 344 ATGACGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 285

Qy 181 ACCTTCACTGCCACACCCATATGCTCCCGGCACTGCTATTAACCTGCGACTCTTTGCATG 240
Db 284 ACCTTCACTGCCACACCCATATGCTCCCGGCACTGCTATTAACCTGCGACTCTTTGCATG 225

Qy 241 CATGCAAACTACTATTATGACAGGGAATGATTAACTAGTCTCTGGAGGACTT 300
Db 224 CATGCAAACTACTATTATGTTTATGGAATGATTAACTAGTCTCTGGAGGACTT 165

Qy 301 GGAGCCACTGTCTGTTGGACTTACTTACCCCATACCATGATGCTGATGGGGTGGAAATT 360
Db 164 GGAGCCACTGTCTGTTGGACTTACTTACCCCATACCATGATGCTGATGGGGTGGAAATT 105

Qy 361 CAAGGTGAGGCAAGAGAAAAACAAGTAAGGAGCAATCTCCCACTGACCCGGGGACAT 420
Db 104 CAAGATCAGGCAAGAGAAAAACAGTAAAGGAAGTAATCTCCCACTGACCCCTGGTACAT 45

Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTC 453
Db 44 AGCACCCCTAGACCCCTACAAAGGACTAATCCCC 12

```

```

US-10-027-632-322492/c
; Sequence 322492, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 322492
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(570)
; OTHER INFORMATION: n = A,T,C or G
; US-10-027-632-322492

```

```

Query Match      25.5%; Score 377.8; DB 17; Length 570;
Best Local Similarity 90.7%; Pred. No. 3e-105;
Matches 411; Conservative 2; Mismatches 39; Indels 1; Gaps 1;

Qy 1 ATGGCCCTCCCTATCATCTTTCTCTTACTGTTCTCTTACCCCTTTTCGCTCTCACT 60
Db 463 ATAGCCCTCCCTATCATATTTTCTCTTACCATTCTCTTACCTGCTTTCACTCTCACT 404

Qy 61 GCACCCCTCCATGCTGCTACACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
Db 403 TCACCTACTCCATGCTGTTGATGACCAAGTAGCTCCCTTACCAAGAGTTTCT-TGGAGA 345

Qy 121 ACGCGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 180
Db 344 ATGACGCTTCTGGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAATCC 285

Qy 181 ACCTTCACTGCCACACCCATATGCTCCCGGCACTGCTATTAACCTGCGACTCTTTGCATG 240
Db 284 ACCTTCACTGCCACACCCATATGCTCCCGGCACTGCTATTAACCTGCGACTCTTTGCATG 225

Qy 241 CATGCAAACTACTATTATGACAGGGAATGATTAACTAGTCTCTGGAGGACTT 300
Db 224 CATGCAAACTACTATTATGTTTATGGAATGATTAACTAGTCTCTGGAGGACTT 165

Qy 301 GGAGCCACTGTCTGTTGGACTTACTTACCCCATACCATGATGCTGATGGGGTGGAAATT 360
Db 164 GGAGCCACTGTCTGTTGGACTTACTTACCCCATACCATGATGCTGATGGGGTGGAAATT 105

Qy 361 CAAGGTGAGGCAAGAGAAAAACAAGTAAGGAGCAATCTCCCACTGACCCGGGGACAT 420
Db 104 CAAGATCAGGCAAGAGAAAAACAGTAAAGGAAGTAATCTCCCACTGACCCCTGGTACAT 45

Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTC 453
Db 44 AGCACCCCTAGACCCCTACAAAGGACTAATCCCC 12

```

Fri Feb 25 16:26:32 2005

```

RESULT 26
US-10-029-386-4312/c
; Sequence 4312, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4312
; LENGTH: 521
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MAP TO AC021619.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: EST HUMAN HIT: BE734284.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUATE 4.00e-26
; OTHER INFORMATION: NT HIT: AF208161.1, EVALUATE 0.00e+00
; US-10-029-386-4312

Query Match      24.6%; Score 364; DB 16; Length 521;
Best Local Similarity 91.7%; Pred. No. 5.2e-101;
Matches 385; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      1062 AAATGTGTGACATGGAACAGGTCACTGACTCCCTGGTCACTTGGCAAGATCAACTTAATC 1121
DB      521 AAATGTGTGACATGGAATGGTGCACCGACTCCCTGGTCACTTGGCAAGATCAACTTAATC 462

QY      1122 CCTAGCAGCAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTAAACGCCAAAGAGG 1181
DB      461 CTTAGCAGCAGTGTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACTGCGCAAGAGG 402

QY      1182 GGGAACTGTTTATTTTAGGAGAAGACGCTGTTATTATGTTAATCAATCCAGAAATGT 1241
DB      401 GGGAACTGTTTATTTTAGGAGAAGATGCTGTTATTACGTTAATCAATCTGGAATCGT 342

QY      1242 CACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGGCTTCAAAA 1301
DB      341 CACCGAAAAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAGGCTTCAAAA 282

QY      1302 CACCGAACGCTGGGCTCTCTCAGCAATGGATGGCTGGTCTCTCCCTTCTTAGGACC 1361
DB      281 CACCGAACCTGGGACCTCTCTCAGCAATGGATGGATGGCTGGATCTCCCTTCTTAGGACC 222

QY      1362 TCTAGCAGCTCTAATATGTTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTTAA 1421
DB      221 TCTAGCAGCTATAATCTGTTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTTAA 162

QY      1422 GTTTGTCTCTTCAGAAATTCGAGCTGTAAAGCTACAGATGGCTTTCAAATGGAACCCCA 1481
DB      161 GTTTGTCTCTTCAGAAATTCGAGCTGTAAAGCTACAAATGGCTCTTCAAAATGGAGCCCA 102

RESULT 27
US-10-027-632-322574
; Sequence 322574, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

```

```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322574
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(551)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-322574

Query Match      23.8%; Score 353; DB 13; Length 551;
Best Local Similarity 92.0%; Pred. No. 1.3e-97;
Matches 412; Conservative 2; Mismatches 30; Indels 4; Gaps 4;

QY      1   ATGGCCCTCCCTTATCATATCTTTTCTCTTTACTCTTCTTACCCCTTTCGCTCTCACT 60
DB      102 ATGGCCCTCCCTTATCATATTTTCTCTTTACTCTTCTTACCCCTTTCGCTCTCACT 161

QY      61   GCACCCCTCCATCTGCTGTACAAACAGTAGTCCCTTACCAAGAGTTTCTATGAAGA 120
DB      162 GCACCCGCTCATGCTGCTGTACCAACAGTAGTCCCTTACCAAGAGTTTCTATGAGA 221

QY      121   ACGGGCTTCTCGAAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAACCTCC 180
DB      222 ATGGGGCTTCTCGAATATTGATGCCCATCATATAGGAGTTTATCTAAGGGAAC-CC 280

QY      181   ACCTTACTGCGCACACCCCATATGCGCGGAACTGCTATACTGCGCACTTTTGCAATG 240
DB      281 ACCTTACTGCGCACACCCCATATGCGCGAATAATGCTGTAACTTGCACACTTTTGCAATG 340

QY      241   CATGCAATATCTTATTGACAGGGAATAATGATTAATCCTAGTTGCTCTGGAGGACTT 300
DB      341 CATGCAATATCTTATTGACAGGGAATAATGATTAATCCTAGTTGCTCTGGAGGACTT 400

QY      301   GGAGCACTGTCTGTGGACTT-ACCTCACCCATACCAAGTATGCTGATGGGGTGGAAAT 359
DB      401 GGAGCACTGTCTGTGGACTTAACTTCACTCATACCGGTATGCTGATGGGGTGGAGT 460

QY      360   TCAAGGTGAGCAAGAGAAAACAAAGTAAGAGGAGCAATCTCCCAACTGACCCCGGACA 419
DB      461 TC-AGATGAGGCAAGAG-AAAACACGTANAGGAAGTAAATCTCCANCTGACNGTGTACA 518

QY      420   TAGCACCCCTAGCCCTACAAAGGACTA 447
DB      519 TAGCACCCCTAACCCCTACAAAGGACTA 546

RESULT 28
US-10-027-632-322575
; Sequence 322575, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

```


ug-09-319-156b-9.rnpb

```

: APPLICANT: ABRURATANI, HIROYUKI
:
: TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
:
: FILE REFERENCE: 084335/166
:
: CURRENT APPLICATION NUMBER: US/10/292,798
:
: CURRENT FILING DATE: 2002-11-13
:
: PRIOR APPLICATION NUMBER: 10/017,161
:
: PRIOR FILING DATE: 2001-12-18
:
: PRIOR APPLICATION NUMBER: JP 2001-246789
:
: PRIOR FILING DATE: 2001-06-18
:
: NUMBER OF SEQ ID NOS: 2070
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 1393
:
: LENGTH: 822900

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1) .. (822900)
FEATURE:
NAME/KEY: CDS
LOCATION: (201) .. (1068)
FEATURE:
NAME/KEY: CDS
LOCATION: (140545) .. (140693)
FEATURE:
NAME/KEY: CDS
LOCATION: (261786) .. (261845)

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1)..(822900)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(1068)
FEATURE:
NAME/KEY: CDS
LOCATION: (140545)..(140693)
FEATURE:
NAME/KEY: CDS
LOCATION: (261786)..(261845)
FEATURE:
NAME/KEY: CDS
LOCATION: (273663)..(273702)
FEATURE:

```

```

, NAME/KEY: CDS
, LOCATION: (347633)..(347711)
, FEATURE:
, NAME/KEY: CDS
, LOCATION: (482589)..(482596)
, FEATURE:
, NAME/KEY: CDS
, LOCATION: (534176)..(534210)
, FEATURE:
, NAME/KEY: CDS
, LOCATION: (822485)..(822700)
, FEATURE:
, NAME/KEY: modified_base
, LOCATION: (4848)..(4947)
, OTHER INFORMATION: a, t, c, g, unknown or other
, FEATURE:
, NAME/KEY: modified_base
, LOCATION: (4966)..(4966)
, OTHER INFORMATION: a, t, c, g, unknown or other
, FEATURE:
, NAME/KEY: modified_base
, LOCATION: (17785)..(17884)
, OTHER INFORMATION: a, t, c, g, unknown or other
, FEATURE:
, NAME/KEY: modified_base
, LOCATION: (61159)..(61258)
, OTHER INFORMATION: a, t, c, g, unknown or other
, FEATURE:
, NAME/KEY: modified_base
, LOCATION: (67605)..(67704)
, OTHER INFORMATION: a, t, c, g, unknown or other
, FEATURE:
, NAME/KEY: modified_base

```

```

1 LOCATION: (367573)..(367573)
2 OTHER INFORMATION: a, t, c, g, unknown or other
3 FEATURE:
4 NAME/KEY: modified_base
5 LOCATION: (367588)..(367588)
6 OTHER INFORMATION: a, t, c, g, unknown or other
7 FEATURE:
8 NAME/KEY: modified_base
9 LOCATION: (367595)..(367595)
10 OTHER INFORMATION: a, t, c, g, unknown or other
11 FEATURE:
12 NAME/KEY: modified_base
13 LOCATION: (367613)..(367614)
14 OTHER INFORMATION: a, t, c, g, unknown or other
15 FEATURE:
16 NAME/KEY: modified_base
17 LOCATION: (367985)..(367985)
18 OTHER INFORMATION: a, t, c, g, unknown or other
19 FEATURE:
20 NAME/KEY: modified_base
21 LOCATION: (367993)..(367993)
22 OTHER INFORMATION: a, t, c, g, unknown or other
23 FEATURE:
24 NAME/KEY: modified_base
25 LOCATION: (367995)..(367996)
26 OTHER INFORMATION: a, t, c, g, unknown or other
27 FEATURE:
28 NAME/KEY: modified_base
29 LOCATION: (367998)..(367998)
30 OTHER INFORMATION: a, t, c, g, unknown or other
31 FEATURE:
32 NAME/KEY: modified_base
33 LOCATION: (370273)..(370277)
34 OTHER INFORMATION: a, t, c, g, unknown or other
35 FEATURE:
36 NAME/KEY: modified_base
37 LOCATION: (370279)..(370279)
38 OTHER INFORMATION: a, t, c, g, unknown or other
39 FEATURE:
40 NAME/KEY: modified_base
41 LOCATION: (370281)..(370282)
42 OTHER INFORMATION: a, t, c, g, unknown or other
43 FEATURE:
44 NAME/KEY: modified_base
45 LOCATION: (370285)..(370287)
46 OTHER INFORMATION: a, t, c, g, unknown or other
47 FEATURE:
48 NAME/KEY: modified_base
49 LOCATION: (370289)..(370291)
50 OTHER INFORMATION: a, t, c, g, unknown or other
51 FEATURE:
52 NAME/KEY: modified_base
53 LOCATION: (385643)..(385742)
54 OTHER INFORMATION: a, t, c, g, unknown or other
55 FEATURE:
56 NAME/KEY: modified_base
57 LOCATION: (409961)..(410060)
58 OTHER INFORMATION: a, t, c, g, unknown or other
59 FEATURE:
60 NAME/KEY: modified_base
61 LOCATION: (410096)..(410096)
62 OTHER INFORMATION: a, t, c, g, unknown or other
63 FEATURE:
64 NAME/KEY: modified_base
65 LOCATION: (417384)..(417483)
66 OTHER INFORMATION: a, t, c, g, unknown or other
67 FEATURE:
68 NAME/KEY: modified_base
69 LOCATION: (724960)..(725059)
70 OTHER INFORMATION: a, t, c, g, unknown or other
71 FEATURE:
72 NAME/KEY: modified_base
73 LOCATION: (726106)..(726205)

```

QY 1073 TCGAACGCTCACTGACCTCCCTGGTCACTTGCAGATCAACTTAACCTCCCTAGCAGCAG 1132
DB 56723 TCGAATGAATTGCCAACTCCCTAGTGCCTTACAAGCCAGCTTAATTCTCTAGCTGCAG 56664
QY 1133 TAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCCAAAAGAGGGGGAACCTGTT 1192
DB 56663 TAGTCTTCAAAACCGGAGAGCCCTAGACTTATTAAACAGCTTAAAGAGGAGGAACCTGCC 56604
QY 1193 TATTTTGGAGAGAACGCTGTTATTTATTTAATCAATCCAGATTGTCTACTGAGAAG 1252
DB 56603 TCTTCTTAGGAAAGAAATGTTGCTATTTTGTGTTAAACAGCTCAGGAATCAATTACTGAAAGG 56544
QY 1253 TTAAGAAATTCGAGATCGAATACAATGTAGAGCAGAGAGCTTCAAAACACCGAACGCT 1312
DB 56543 TCAAGAAATTAAGAGAACAAATAGAAAGTAGAAAGAGAGCTTGAACACTCAGGGCGCT 56484
QY 1313 GGGGCGCTCTCAGCCAAATGGATGCCCTCGGTTCTCCCTTCTTAGGACCTCTAGCAGCTC 1372
DB 56483 TGAATATGTTTAAACCAATGGATACCTCGGCTCCTCCCTTCTAGGCGCTGCGACAGC-- 56426
QY 1373 TAATATGTTACTCTCTTTGGACCGCTGTATCTTTTAACTCCTTGTAAAGTTTGTCTCT 1432
DB 56425 ----CATCTACTCGTTTTTGGGCGCTTGCAATTTTAACTCCTCTGTCAAAATTTGTTCT 56370
QY 1433 CCAGAAATTCAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
DB 56369 -CAGGATCGAGGCCATCAAGGTACAAATGCTTACAAATGGAACCTCA 56322

RESULT 32
US-08-979-847-108
; Sequence 108, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACT
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1329 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-979-847-108
Query Match 20.4%; Score 302; DB 8; Length 1329;
Best Local Similarity 92.7%; Pred. No. 1.1e-81;
Matches 317; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1140 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCCAAAAGAGGGGGAACCTGTTATTTT 1199
DB 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCCAAAAGAGGGGGAACCTGTTATTTT 60
QY 1200 AGGAGAAAGCGCTGTTATTATGTTAATCAATCCAGAAATGTCTCACTGAGAAAGTTAAAGA 1259
DB 61 AGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTACTGGAAGTTAAAGA 120
QY 1260 AATTGAGATCGAATAACAATGTAGAGCAGAGGAGCTTCAAAACACCGAACCGCTGGGCGCT 1319
DB 121 AATTGAGATCGAATAATAATGTAGAGCAGAGGAGCTTCAAAACACTGCACCTGGGCGCT 180
QY 1320 CCTCAGCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGCAGCTCTAAATTT 1379
DB 181 CCTCAGCAATGGATGCCCTGGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
QY 1380 GTTACTCTCTTTGGACCGCTGTATCTTTAACTCCTTGTAAAGTTTGTCTCTCCAGAAAT 1439
DB 241 TTTACTCTCTTTGGACCGCTGTATCTTCACTTCTTGTAAAGTTTGTCTCTCCAGAAAT 300
QY 1440 TGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGAACCCCA 1481
DB 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCCA 342

RESULT 33
US-10-114-104-108
; Sequence 108, Application US/10114104
; Publication No. US20030198647A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; STREET: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,104
; FILING DATE: 03-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/979,847
; FILING DATE: 26-NOV-1997
; ATTORNEY/AGENT INFORMATION:

; PRIOR APPLICATION NUMBER: PCT/US01/00664					
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	PCT/US01/00669				
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	PCT/US01/00665				
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	PCT/US01/00668				
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	PCT/US01/00663				
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	PCT/US01/00662				
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	PCT/US01/00661				
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	PCT/US01/00670				
; PRIOR FILING DATE:	2001-01-30				
; PRIOR APPLICATION NUMBER:	US 60/234,687				
; PRIOR FILING DATE:	2000-09-21				
; PRIOR APPLICATION NUMBER:	US 09/608,408				
; PRIOR FILING DATE:	2000-06-30				
; PRIOR APPLICATION NUMBER:	US 09/774,203				
; PRIOR FILING DATE:	2001-01-29				
NUMBER OF SEQ ID NOS: 49117					
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1					
SEQ ID NO 14030					
LENGTH: 426					
TYPE: DNA					
ORGANISM: Homo sapiens					
FEATURE:					
OTHER INFORMATION: MAP TO AL022067.1					
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.95					
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1					
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3					
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2					
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.1					
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6					
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2					
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1					
US-09-864-761-14030					
Query Match 20.0%; Score 296; DB 9; Length 426;					
Best Local Similarity 85.7%; Pred No. 4.2e-80;					
Matches 365; Conservative 0; Mismatches 30; Indels 31; Gaps 2;					
Qy	1087	GACTCCCTGGTCACCCTTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTTCAAAT	1146		
Dd	1	GACTCCCTGGTCACCCTTGCAAGATCAACTTAACTCTAGCATCAGTAGTCTTTCAAAT	60		
Qy	1147	CGAAGAGCTTTGAGCTTGCTACCGCCAAAAGAGGGGAACCTGTATTATTTTAGG----	1202		
Dd	61	TGAAGAGCTTTGAGCTTGCTAACCTCTGAAAGAGGGGGAAGCTGTATTATTTTAGGGGAA	120		
Qy	1203	-----AGAAGAACGCCTGTTATTATCTTAATCAATCAATCCAGA	1236		
Dd	121	GAATGTTGTTATTATGTTATTTTACGCCGAAGATGTTGTTATTATGTTAATCAATCCTGA	180		
Qy	1237	ATTGTCAC TGAGAAA GTTAAAGAAA TCGAGATCGAATA CAATGTAGAGCAGAGGAGCTT	1296		
Dd	181	ATTGTCACAGAGAAAGTTGAGAAA TCGAGATTGAATACAACG TAGNACAGAGGAGCTT	240		
Qy	1297	C-AAACACCGAA CGCTGGGGCCCTCTCAGCCAATGGATGCCCTGGGTTCTCCCCCTCTT	1355		
Dd	241	CAAAAAACACGAG ACCCTGGGGCCCTCTCAGCCAAATGGATGCCCTGGATTTCTCCCCTTCTT	300		
Qy	1356	AGACCTCTAGCAGCTCTAATATTGTTACTTCTTTGGACCCCTGTATCTTTAACTTCCT	1415		
Dd	301	AGGATCTCTAGCAGCTCTAATATTGATCTCTCTTTGGACCCCTGTATCTTTAACTTCCT	360		
Qy	1416	TGTTAAGTTTGTCCTCTTCAGAAATGAAGCTGTAAAGCTACAGATGGTCTTACAAATGGA	1475		
Dd	361	TGTTAAGTTTGTCCTCTTCAGAAATGAAGCTGTAAAGCTACAAATCGTCTTCAAATGGA	420		
Qy	1476	ACCCCA	1481		


```
Db          421 ACCCCA 426
|||||
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316412
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; FEATURE:
; LOCATION: (1)..(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-316412

Query Match      15.5%; Score 229.8; DB 13; Length 625;
Best Local Similarity 88.7%; Pred. No. 1.3e-59;
Matches 260; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 317 GGACTTACTTACCCNATACAGTATGCTGTGATGGGGTGGGAATTCAGGTCAAGCAAGAG 376
Db 625 GAATTCACCTTACCCATACCGGTATGCTGTGATGGGGTGGAGTTCAAGATCAGGCAAGAG 566
QY 377 AAAACAAGTAAAGGAAGCAATCTCCCACTGACCGGGGACATAGCACCCCTAGCCCT 436
Db 565 AAAACATGTAAAGGAAGTATCTCCCACTGACTTGGGTACATTAGCACCCCTAGCCCT 506
QY 437 ACAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACCCATCTCGCCTGGTGA 496
Db 505 GCAAGGACTAGATCTCTCAAACTACATGAACCCCTCCATACCCCATCTACCTGGTAA 446
QY 497 GCCTATTTAATACCACTTCTCGGCTCCATGAGGTCTCAGCCCAAAACCTACTAACT 556
Db 445 GCCTATTTAATACCACTTCTCGGCTCCATGAGGTCTCAG-AGAAAACCCCTACTAACT 387
QY 557 GTTGATGTGCTTCCCTCGCTTCCAGGCCATACATTTCAATCCCTGTCTCT 609
Db 386 GTTGATGTGCTTCCCTCGCTTCCAGGCCATACATTTCAATCCCTGTATCT 334

RESULT 36
US-10-027-632-316412/c
; Sequence 316412, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 316412
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; FEATURE:
; LOCATION: (1)..(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-316412

Query Match      15.5%; Score 229.8; DB 17; Length 625;
Best Local Similarity 88.7%; Pred. No. 1.3e-59;
Matches 260; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 317 GGACTTACTTACCCATACAGTATGCTGTGATGGGGTGGGAATTCAGGTCAAGCAAGAG 376
Db 625 GAATTCACCTTACCCATACCGGTATGCTGTGATGGGGTGGAGTTCAAGATCAGGCAAGAG 566
QY 377 AAAACAAGTAAAGGAAGCAATCTCCCACTGACCGGGGACATAGCACCCCTAGCCCT 436
Db 565 AAAACATGTAAAGGAAGTATCTCCCACTGACTTGGGTACATTAGCACCCCTAGCCCT 506
QY 437 ACAAGGACTAGTTCTCTCAAAACTACATGAACCCCTCCGTACCCATCTCGCCTGGTGA 496
Db 505 GCAAGGACTAGATCTCTCAAACTACATGAACCCCTCCATACCCCATCTACCTGGTAA 446
QY 497 GCCTATTTAATACCACTTCTCGGCTCCATGAGGTCTCAGCCCAAAACCTACTAACT 556
Db 445 GCCTATTTAATACCACTTCTCGGCTCCATGAGGTCTCAG-AGAAAACCCCTACTAACT 387
QY 557 GTTGATGTGCTTCCCTCGCTTCCAGGCCATACATTTCAATCCCTGTCTCT 609
Db 386 GTTGATGTGCTTCCCTCGCTTCCAGGCCATACATTTCAATCCCTGTATCT 334

RESULT 37
US-10-027-632-86541/c
; Sequence 86541, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86541
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-86541

Query Match      15.5%; Score 229.6; DB 13; Length 619;
Best Local Similarity 86.3%; Pred. No. 1.5e-59;
Matches 253; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 317 GGACTTACTTCAACCATACCAAGTATGTCCTGATGGGGTGAATTCAGGTCAAGGTCAGGCAAGAG 376
Db 619 GAATTCACCTTCAACCATACCAAGTATGTCCTGATGGGGTGAATTCAGGTCAAGGTCAGGCAAGAG 560

QY 377 AAAACAAGTAAGGAAGCAATCTCCCAACTGACCCGGGGACATAGCACCCCTAGCCCT 436
Db 559 AAAACAATGTAAGGAAGTCACTCTCCCAACTGACCCGGGGTACATAGCACCCCTAGCCCT 500

QY 437 ACAAGGACTAGTCTCTCAAACTACATGAACCCCTCGTACCCATCTCGCTGTGTA 496
Db 499 GCAAGGACTAGTCTCTCAAACTACATGAACCCCTCGTACCCATCTCGCTGTGTA 440

QY 497 GCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCCCTACTAACT 556
Db 439 GCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCCCTACTAACT 380

QY 557 GTTGGATGTGCTCCCTCGTCACTTCAAGGCAATTCATTCATCCCTGTCTCT 609
Db 379 GTTGGATGTGCTCCCTCGTCACTTCAAGGCAATTCATTCATCCCTGTCTCT 327

RESULT 38
US-10-027-632-86541/c
; Sequence 86541, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86541
; LENGTH: 619
```

```
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-86541

Query Match      15.5%; Score 229.6; DB 17; Length 619;
Best Local Similarity 86.3%; Pred. No. 1.5e-59;
Matches 253; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 317 GGACTTACTTCAACCATACCAAGTATGTCCTGATGGGGTGAATTCAGGTCAAGGTCAGGCAAGAG 376
Db 619 GAATTCACCTTCAACCATACCAAGTATGTCCTGATGGGGTGAATTCAGGTCAAGGTCAGGCAAGAG 560

QY 377 AAAACAAGTAAGGAAGCAATCTCCCAACTGACCCGGGGACATAGCACCCCTAGCCCT 436
Db 559 AAAACAATGTAAGGAAGTCACTCTCCCAACTGACCCGGGGTACATAGCACCCCTAGCCCT 500

QY 437 ACAAGGACTAGTCTCTCAAACTACATGAACCCCTCGTACCCATCTCGCTGTGTA 496
Db 499 GCAAGGACTAGTCTCTCAAACTACATGAACCCCTCGTACCCATCTCGCTGTGTA 440

QY 497 GCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCCCTACTAACT 556
Db 439 GCCTATTATTAACCAACCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCCCTACTAACT 380

QY 557 GTTGGATGTGCTCCCTCGTCACTTCAAGGCAATTCATTCATCCCTGTCTCT 609
Db 379 GTTGGATGTGCTCCCTCGTCACTTCAAGGCAATTCATTCATCCCTGTCTCT 327

RESULT 39
US-09-864-761-7501/c
; Sequence 7501, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7501
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL139245.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; US-09-864-761-7501

Query Match      14.8%; Score 219.4; DB 9; Length 559;
Best Local Similarity 91.7%; Pred. No. 2e-56;
Matches 232; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      1229 AATCCAGAAATGTCACAGAGAGTAAAGAAATTCGAGATCGAATAAATGTAGAGCAG 1288
DB
QY      1289 AGGAGCTTCAAAACACCGAACGCTGGGGCTCTCAGCAATGGATGCCCTGGTTCTCC 1348
DB
QY      1349 CCTTCTTAGGACCTCTAGCAGCTTAATATTGTTACTCTCTTGGACCTCTGTATCTTTA 1408
DB
QY      1409 ACCTCTTGTAAAGTTGTCTCTCCAGAAATTAAGCTGTAAGCTACAGATGGTCTTAC 1468
DB
QY      1469 AATGGAACCCCA 1481
DB      319 AATGGAGCCCCA 307

RESULT 40
US-10-029-386-20259
; Sequence 20259, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20259
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; OTHER INFORMATION: MAP TO AP001697.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.91
; OTHER INFORMATION: NT HIT: AF208161.1, EVALUE 1.00e-108
; OTHER INFORMATION: SWISSPROT HIT: P28808, EVALUE 6.00e-12
; OTHER INFORMATION: EST_HUMAN HIT: BE734284.1, EVALUE 1.00e-108
; US-10-029-386-20259

Query Match      14.4%; Score 213; DB 16; Length 494;
Best Local Similarity 85.6%; Pred. No. 1.7e-54;
Matches 261; Conservative 0; Mismatches 40; Indels 4; Gaps 2;

QY      1106 AAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGACTTGC 1165
DB      193 AAGGACTTCAAAAGCGGTGTATGTCAGTAGTCCTTCAAAATCGAAGAGCTTTAGAAATTGC 252
QY      1166 TAACCGCCAAAAGAGGGGGAACCTGTTTATTTTAGGAGAAGACGCTGTTATTATGTTA 1225
DB      253 TAATCACTGAGAGAGGGGAACGTTTTTATTTTAGGGGAAGAAATGCTGTTATTATGTTA 312
QY      1226 ATCAATCCAGAAATGTCACCTGAGAGAAAGTTAAAGAAATTCGAGATCGAATCAATGTAGAG 1285
DB      313 ATCAATTCGGAAATCATCACCAAGAAAGTTAAAGAAATTCAGATCGAATCAACGTAGAA 372
QY      1286 CAGAGGAGCTT-CAAAACACCGAACGCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGGTT 1344
DB      373 CAGAGGAGCTTAAAAAACACTGGACCTGGGGCTCTCTCAGCCAAATGGATGCCCTGGATT 432
QY      1345 CTCCCTTCTTAGGACCTCTAGCAGCTTAATATTGTTACTCTCTTTGGACCTGTATC 1404
DB      433 CTCCCTTCTTAGGACCTCTAGCAGCT---ATATTCTACTCTCTTTGGACCTGTATC 489
QY      1405 TTTAA 1409
DB      490 TTTAA 494

Search completed: February 21, 2005, 19:18:14
Job time : 827.526 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:52:54 ; Search time 4605.5 Seconds
(without alignments)
12240.400 Million cell updates/sec

Title: US-09-319-156B-9
Perfect score: 1481
Sequence: 1 atggccctccctatcatcac.....gtcttacaagtgaacccca 1481

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1333.8	90.1	2500	3	CR622175 full-length
2	1333.8	90.1	2716	3	CR613169 full-length
3	1333.8	90.1	2748	3	CR605851 full-length
4	1333.8	90.1	2748	3	CR625046 full-length
5	1333.8	90.1	2749	3	CR617248 full-length
6	839	56.7	935	5	CR6191741 full-length
7	813.6	54.9	955	5	CR617952 full-length
8	752.2	50.8	931	5	CR619656 full-length
9	747.2	50.0	872	5	CR626647 full-length
10	741	50.0	903	5	CR608752 full-length
11	737.2	49.8	883	5	CR608733 full-length
12	728.4	49.2	877	5	CR617619 full-length
13	714	48.2	836	5	CR618766 full-length
14	709.8	47.9	921	5	CR608734 full-length
15	705.8	47.7	844	5	CR617111 full-length
16	703.2	47.5	828	5	CR617907 full-length
17	698.6	47.2	846	5	CR6168078 full-length
18	696.2	47.0	995	5	CR6189657 full-length
19	695	46.9	924	5	CR617314 full-length
20	647	43.7	723	1	AUI138405 full-length
21	629.4	42.5	631	8	AQ261133 full-length
22	628.6	42.4	702	1	AUI138097 full-length
23	623.8	42.1	924	5	CR609328 full-length
24	609.8	41.2	714	2	BE734284 full-length

C 25	580.4	39.2	658	5	CR622175	full-length cDNA clone CS0D1051YM13 of Placenta Cot 25-normalized
C 26	555.8	37.5	792	5	CR622175	of Homo sapiens (human)
C 27	537.6	36.3	814	5	CR622175	HTC; CNSIT_cDNA
C 28	536.4	36.2	586	4	CR622175	Homo sapiens (human)
C 29	535.2	36.1	790	4	CR622175	Homo sapiens
C 30	478.6	32.3	760	5	CR622175	Homo sapiens
C 31	475.4	32.1	554	7	CR622175	1 (bases 1 to 2500)
C 32	444	30.0	1071	5	CR622175	Location/Qualifiers
C 33	421.6	28.5	944	5	CR622175	/organism="Homo sapiens"
C 34	416.4	28.1	890	5	CR622175	/mol_type="mRNA"
C 35	412.2	27.8	998	5	CR622175	/db_xref="taxon:9606"
C 36	412.2	27.8	1058	5	CR622175	/clone="CS0D1051YM13"
C 37	411.4	27.8	905	5	CR622175	/tissue type="Placenta Cot 25-normalized"
C 38	406.4	27.4	564	5	CR622175	/plasmid="PCVSPORT_6"
C 39	402	27.1	1019	5	CR622175	
C 40	401.4	27.1	484	5	CR622175	
C 41	399.6	27.0	494	5	CR622175	
C 42	390.2	26.3	427	1	CR622175	
C 43	389.8	26.3	718	1	CR622175	
C 44	373.2	25.2	966	5	CR622175	
C 45	371.4	25.1	586	8	CR622175	

ALIGNMENTS

RESULT 1	CR622175	2500 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR622175	full-length cDNA clone CS0D1051YM13 of Placenta Cot 25-normalized			
DEFINITION	CR622175	of Homo sapiens (human)			
ACCESSION	CR622175	GI:50502982			
VERSION	CR622175.1	GI:50502982			
KEYWORDS	HTC; CNSIT_cDNA				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2500)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Paraday Avenue				
REFERENCE	2 (bases 1 to 2500)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr				
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the PCVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..2500				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CS0D1051YM13"				
	/tissue type="Placenta Cot 25-normalized"				
	/plasmid="PCVSPORT_6"				
ORIGIN					
Query Match	90.1%;	Score 1333.8;	DB 3;	Length 2500;	
Best Local Similarity	93.8%;	Pred. No. 0;			
Matches 1389;	Conservative	Mismatches 92;	Indels 0;	Gaps 0;	
QY	1	ATGCCCTCCCTATCATCATTTTCTTTACTGTTCTTACCCCTTTCCTACT	60		
Db	530	ATGCCCTCCCTATCATCATTTTCTTTACTGTTCTTCTTACTGTTCTTCTTCTACT	589		

QY 61 GCACCCCTCCATGCTGCTGTAACAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 120
 DB 590 GCACCCCTCCATGCTGCTGTAACAACAGTAGCTCCCTTACCAAGAGTTTCTATGAAGA 649
 QY 121 AGCGGCTTCTGGAAATATTGATGCCCATCATAGAGTTTATCTAAGGGAACCTCC 180
 DB 650 ATGACGCTCCGGAATATTGATGCCCATCATAGAGTTTATCTAAGGGAACCTCC 709
 QY 181 ACCTTCACTGCCACACCATATGCTCCGCAACTGCTATACTTGCACCTCTTTGCATG 240
 DB 710 ACCTTCACTGCCACACCATATGCTCCGCAACTGCTATACTTGCACCTCTTTGCATG 769
 QY 241 CATGMAATATCTATTATGGACAGGAAATGATTAATCTAGTGTCTTGGAGGACTT 300
 DB 770 CATGMAATATCTATTATGGACAGGAAATGATTAATCTAGTGTCTTGGAGGACTT 829
 QY 301 SGAGCCACTGCTGTGGACTTACTTCAACCATACAGTATGCTCTGATGGGGTGAAT 360
 DB 830 GGAGTCACTGTCTGTGGACTTACTTCAACCATACAGTATGCTCTGATGGGGTGAAT 889
 QY 361 CAAGGTGAGGCAAGGAAACAAAGTAAAGGAAGCAATCTCCAACTGACCGGGGACAT 420
 DB 890 CAAGATCAGCAAGGAAACAAAGTAAAGGAAGTAACTCCCAACTCACCGGGTACAT 949
 QY 421 AGACCCCTAGCCCTPACAAAGACTAGTTCTCTCAAACTACATGAACCTCCGTACC 480
 DB 950 GGACCTCTAGCCCTPACAAAGACTAGATCTCTCAAACTACATGAACCTCCGTACC 1009
 QY 481 CATACTCGCTGTGAGCTTATTAATACACCTCACCTCGGCTCCATGAGTCTCAGCC 540
 DB 1010 CATACTCGCTGTGAGCTTATTAATACACCTCACCTCGGCTCCATGAGTCTCAGCC 1069
 QY 541 CAAACCCCTACTAATCTGATGCTCCCTCGCTTCCAGGCTACATCAATC 600
 DB 1070 CAAACCCCTACTAATCTGATGCTCCCTCGCTTCCAGGCTACATCAATC 1129
 QY 601 CTGTGCTGCAATGGAACAACTTACGACAGAAATAAACCACTTCCGTTTAGTA 660
 DB 1130 CTGTGCTGCAATGGAACAACTTACGACAGAAATAAACCACTTCCGTTTAGTA 1189
 QY 661 GGACCTCTGTTTCCAACTGGAATAAACCACTTCCAACTCACCTGTGTAAATTT 720
 DB 1190 GGACCTCTGTTTCCAACTGGAATAAACCACTTCCAACTCACCTGTGTAAATTT 1249
 QY 721 AGCAATCTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAACTCCCAACGA 780
 DB 1250 AGCAATCTATAGACACCAACAGCTCCCAATGATCAGGTGGGTAACTCCCAACGA 1309
 QY 781 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGTTACCTCAGCCTATCATTTTGG 840
 DB 1310 ATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGTTACCTCAGCCTATCATTTTGG 1369
 QY 841 AATGGCTCTTCAGAACTATGCTCTCTCAATCTTAGTGGCCCTATGACCACTAC 900
 DB 1370 AATGGCTCTTCAGAACTATGCTCTCTCAATCTTAGTGGCCCTATGACCACTAC 1429
 QY 901 ACTGAACAAGATTATACAAATCATGCTACCTTAAGCCCAACAAAGAGTACCATT 960
 DB 1430 ACTGAACAAGATTATACAAATCATGCTACCTTAAGCCCAACAAAGAGTACCATT 1489
 QY 961 CTTTCTTTTGTATCAGACAGAGTGTAGGAGACTAGGTACTGGCATTTGGCAGTATC 1020
 DB 1490 CTTTCTTTTGTATGAGGACAGAGTGTAGGAGACTAGGTACTGGCATTTGGCAGTATC 1549
 QY 1021 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTACATGGAACAG 1080
 DB 1550 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGGTACATGGAACAG 1609
 QY 1081 GTCACTGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1140
 DB 1610 GTGCCGACTCCCTGGTCACTTTGCAAGATCAACTTAACTCCCTAGCAGAGTAGTCCCT 1669

QY 1141 CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGAAACCTGTTTATTTTA 1200
 DB 1670 CAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAAACCTGTTTATTTTA 1729
 QY 1201 GGAGAAGAACGCTGTTTATTATGTTTAAATCAATCCAGAAATGCTCACTGAGAAAGTTAAAGAA 1260
 DB 1730 GGGAAGAAATGCTGTTTATTATGTTTAAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 1789
 QY 1261 ATTGAGATCGAATACAAATCTAGAGAGAGAGCTTCAAAACACCGAACCTCGGGGCTC 1320
 DB 1790 ATTGAGATCGAATACAAATCTAGAGAGAGAGCTTCAAAACACCTCGGGGCTC 1849
 QY 1321 CTGAGCAATGAGTGGCTCGGCTTCCCTTCTTAGGAGCTCTAGCAGCTCTAAATATTG 1380
 DB 1850 CTGAGCAATGAGTGGCTCGGCTTCCCTTCTTAGGAGCTCTAGCAGCTCTAAATATTG 1909
 QY 1381 TTACTCTCTTTGAGCCCTGTATCTTTAACTCTCTTGTAAAGTTGCTCTTCCAGAAAT 1440
 DB 1910 CTACTCTCTTTGAGCCCTGTATCTTTAACTCTCTTGTAAAGTTGCTCTTCCAGAAATC 1969
 QY 1441 GAAGCTGTAAGCTACAGATGCTTTACAAATGGAACCCCA 1481
 DB 1970 GAAGCTGTAAGCTACAAATGGAACCCCAAGATGCAATGCAAT 2010

RESULT 2
 CR613169 2716 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CSODE013Y120 of Placenta of Homo sapiens
 DEFINITION
 ACCESSION CR613169
 VERSION CR613169.1 GI:50493976
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2716)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 2716)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Genoscope - Centre National de Sequencage : Submitted (20-JUL-2004) EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr BP 191 91006 EVRY cedex - FRANCE)

COMMENT
 - Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.
 FEATURES
 Location/Qualifiers
 1..2716
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODE013Y120"
 /tissue type="Placenta"
 /plasmid="pCMVSPORT_6"

ORIGIN
 Query Match 90.1%; Score 1333.8; DB 3; Length 2716;
 Best Local Similarity 93.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1 ATGGCCCTCCCTTATCATACTTTTCTCTTACTGTCTTACCCCTTTCGCTCTCACT 60
 DB 773 ATGGCCCTCCCTTATCATATTTTCTTACTGTCTTACCCCTTTCCTCACTCTCACT 832

```
Qy 61 GCACCCCTCCATGCTGTGTACCAACAGTAGCTCCCTTACCAGAGTTTCTATGAAGA 120
Db 833 GCACCCCTCCATGCTGTGTATGACAGTAGCTCCCTTACCAGAGTTTCTATGAAGA 892
Qy 121 AGCGGCTTCTCGAAATATTGATGCCCCATCATATAGAGTTTATCTAAGGAAACTCC 180
Db 893 ATGACGGTCCCGAAATATTGATGCCCCATCATATAGAGTTTCTAAGGAACTCC 952
Qy 181 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAAATCTGCCACTCTTTGCATG 240
Db 953 ACCTTCACTGCCACACCCATATGCCCGCAACTGCTATAAATCTGCCACTCTTTGCATG 1012
Qy 241 CATGCAAACTACTCATTTATGAGCAGGAGAAATGATTAATCTAGTTGTCTCGAGGACTT 300
Db 1013 CATGCAAACTACTCATTTATGAGCAGGAGAAATGATTAATCTAGTTGTCTCGAGGACTT 1072
Qy 301 GGAGCCACTGCTGTGGACTTACTTCAACCCATACAGATATGCTGATGGGGTGAAT 360
Db 1073 GGAGTCACTGCTGTGGACTTACTTCAACCAACTGCTATGCTGATGGGGTGGAGTT 1132
Qy 361 CAAGGTCAGCAAGAGAGAAACAAAGTAAGAGCAATCTCCCAACTGACCCCGGGACAT 420
Db 1133 CAAGATCAGGCAAGAGAGAAACATGTAAAGAGTAATCTCCCAACTCACCCGGGTACAT 1192
Qy 421 AGCACCCCTAGCCCTACAAAGGACTAGTTCTCTCAAACTACATGAAACCCCTCCGTACC 480
Db 1193 GGACCTCTAGCCCTACAAAGGACTAGATCTCTCAAACTACATGAAACCCCTCCGTACC 1252
Qy 481 CATACTCGCCTGTGGACCTATTATTAACACCCCTCACTCGGCTCCATGAGGCTCAGCC 540
Db 1253 CATACTCGCCTGTGAAGCCTATTATTAACACCCCTCACTCGGCTCCATGAGGCTCAGCC 1312
Qy 541 CAAAACCTCTACTACTGCTGTGGATGCTGCTCCCTCGACCTCAGGCCATACATTTCAATC 600
Db 1313 CAAAACCTCTACTACTGCTGTGGATGCTGCTCCCTCGAATCTCAGGCCATATGTTTCAATC 1372
Qy 601 CTGTGCTCTGAACTGAAACAACTTCAGACACAGAAATAAACACCACTTCGCTTTTAGTA 660
Db 1373 CTGTGCTCTGAACTGAAACAACTTCAGACACAGAAATAAACACCACTTCGCTTTTAGTA 1432
Qy 661 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCTCACTGTGTAAATTT 720
Db 1433 GGACCTCTGTTTCCAACTCTGGAATAACCCATACCTCAAACTCTCACTGTGTAAATTT 1492
Qy 721 AGCAATACTATAGACACAAACAGCTCCCAATGATGATGATGATGATGATGATGATGAT 780
Db 1493 AGCAATACTATAGACACAAACCTCCCAATGATGATGATGATGATGATGATGATGAT 1552
Qy 781 ATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACTCAGCTCAGCTATCATTTGTTG 840
Db 1553 ATAGTCTGCTTACCCTCAGGAATATTTTGTCTGTGTACTCAGCTCAGCTATCATTTGTTG 1612
Qy 841 AATGGCTCTTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 1613 AATGGCTCTTCAGAACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1672
Qy 901 ACTGAAAGATTTATCAATCATCTGCTAGCTTAAGCCCAACACCAAAAGAGTACCCATT 960
Db 1673 ACTGAAAGATTTATACAGTTATGCTATATCTAATGATGATGATGATGATGATGATGATGAT 1732
Qy 961 CTTCTCTTTGTTATCAGAGCAGGAGTGTAGGAGACTAGTGTAGTGTAGTGTAGTGTAGTGT 1020
Db 1733 CTTCTCTTTGTTATCAGAGCAGGAGTGTAGGAGACTAGTGTAGTGTAGTGTAGTGTAGTGT 1792
Qy 1021 ACAAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTTGATGATGAAACAG 1080
Db 1793 ACAACCTCTACTCAGTTCTACTACAACTATCTCAAGAAATAAATGTTGATGATGAAACAG 1852
Qy 1081 GTCACTGACTCCCTGGTCACTTGCAGATCAACTTAATCTCCCTAGCAGAGTAGTCTCT 1140
Db 1853 GTCCGCGACTCCCTGGTCACTTGCAGATCAACTTAATCTCCCTAGCAGAGTAGTCTCT 1912
Qy 1141 CAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAAAGAGGGGAACTCTGTTTATTTTA 1200
```

```
Db 1913 CAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCTGAAAGAGGGGAAACCTGTTTATTTTA 1972
Qy 1201 GGAGAGAAACGCTGTTTATTTATGTTAATCAATCAGAAATTTGTCACTGAGAAAGTTAAAGAA 1260
Db 1973 GGGAGAAATGCTGTTTATTTATGTTAATCAATCGGAAATGCTCACTGAGAAAGTTAAAGAA 2032
Qy 1261 ATTGAGATCGAATACAATGTTAGAGCAGAGAGCTTCAAAAACACCGAAGCTGGGGCCTC 1320
Db 2033 ATTGAGATCGAATACAATGTTAGAGCAGAGAGCTTCAAAAACACCGAAGCTGGGGCCTC 2092
Qy 1321 CTCAGCAATGATGCCCTGGCTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAAATATTG 1380
Db 2093 CTCAGCAATGATGCCCTGGCTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAAATATTG 2152
Qy 1381 TTACTCTCTTTGGACCCCTGTATCTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGAATT 1440
Db 2153 CTACTCTCTTTGGACCCCTGTATCTTTAAACCTCTTTTAAAGTTTGTCTCTTCCAGAATC 2212
Qy 1441 GAAGCTGTAAAGCTACAGATGCTTTTCAAAATGGAAACCCCA 1481
Db 2213 GAAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCCAA 2253

RESULT 3
LOCUS CR605851
DEFINITION full-length cDNA clone CS0DE012VJ24 of Placenta of Homo sapiens (human) .
ACCESSION CR605851
VERSION CR605851.1 GI:50486658
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2748)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 2748)
REFERENCE 2 (bases 1 to 2748)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
source
1 . 2748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012VJ24"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 90.1%; Score 1333.8; DB 3; Length 2748;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
Qy 1 ATGGCCCTCCCTTATCATCTTTTCTCTTACTTGTCTTCTTACCCCTTTCGCTCTCACT 60
Db 772 ATGGCCCTCCCTTATCATATTTTTCTCTTATCTGTTCTTTTACCTCTTTCACCTCACT 831
Qy 61 GCACCCCTCCATGCTGTGTACCAACAGTAGCTCCCTTACCAGAGTTTCTATGAAGA 120
```

Db 832 GACCCCTCCATGCGCTGTATGACAGTAGTCCCTTACCAAGAGTTCTATGGAGA 891
Qy 121 ACAGCGCTTCTCGAAATATTGATGCCCATCATATAGAGTTTATCTAAAGGAACTCC 180
Db 892 ATGACGGTCCCGAAATATTGATGCCCATCGTATAGAGTCTTCTTAAGGAAACCC 951
Qy 181 ACCTTCACTGCCACACCATATGATGCCCGCACTGCTATCACTGCCACCTTTGCAATG 240
Db 952 ACCTTCACTGCCACACCATATGATGCCCGCACTGCTATCACTGCCACCTTTGCAATG 1011
Qy 241 CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTTAGTTGCTCGAGGACTT 300
Db 1012 CATGCAAACTACTCATTTATGGACAGGAAATGATTAATCTTAGTTGCTCGAGGACTT 1071
Qy 301 GGAGCCACTGTCTGTGGACTTACTTACCCTATACAGATATGCTGTGATGGGGTGAATTT 360
Db 1072 GGAGTCACTGTCTGTGGACTTACTTACCCTATACAGATATGCTGTGATGGGGTGAATTT 1131
Qy 361 CAAGTCAAGCAAGAAACAACTTAAGGAAGCAATCTCCCACTGACCCGGGACAT 420
Db 1132 CAAGATCAGGCAAGAAACAACTTAAGGAAGCAATCTCCCACTGACCCGGGTACAT 1191
Qy 421 AGCACCCCTAGCCCTTACAAAGGACTAGTTCTCTCAAACTACATGAACCCCTCCGTACC 480
Db 1192 GGCACCTCTAGCCCTTACAAAGGACTAGATCTCTCAAACTACATGAACCCCTCCGTACC 1251
Qy 481 CATACTCGCTCGTGTAGCTATTTAATACACCTCACTCGGCTCGATGAGTCTCAGCC 540
Db 1252 CATACTCGCTCGTGTAGCTATTTAATACACCTCACTCGGCTCGATGAGTCTCAGCC 1311
Qy 541 CAAACCCCTACTACTGTTGGATGTCCTCCCTGCACTTCAGGCAATACATTTCAATC 600
Db 1312 CAAACCCCTACTACTGTTGGATGTCCTCCCTGCACTTCAGGCAATACATTTCAATC 1371
Qy 601 CTGTTTCTGAACTTGAACCACTTTCAGCACAGAAATAAACCACTTCCTGTTTAGTA 660
Db 1372 CTGTTACTGAACTTGAACCACTTTCAGCACAGAAATAAACCACTTCCTGTTTAGTA 1431
Qy 661 GGAACCTTGTGTTTCCAACTGGAATAAACCACTTCAACCTCACTGTTGTAATTT 720
Db 1432 GGAACCTTGTGTTTCCAACTGGAATAAACCACTTCAACCTCACTGTTGTAATTT 1491
Qy 721 AGCAATACTATAGACAAACCACTCCCACTGAGTGGTAACTCCACACGA 780
Db 1492 AGCAATACTATAGACAAACCACTCCCACTGAGTGGTAACTCCACACGA 1551
Qy 781 ATAGTCTGCTACCTCAGGAATATTTTGTGTGTGTTACCTCAGCTTATCATTTGTTG 840
Db 1552 ATAGTCTGCTACCTCAGGAATATTTTGTGTGTGTTACCTCAGCTTATCATTTGTTG 1611
Qy 841 AATGGCTCTCAGAACTATGCTTCCCTCACTTCTTAGTGGCCCTATGACCATCTAC 900
Db 1612 AATGGCTCTCAGAACTATGCTTCCCTCACTTCTTAGTGGCCCTATGACCATCTAC 1671
Qy 901 ACTGAACAAGATTATACAACTATGCTGACCTTAAGCCCAACAAAGAGTACCCATT 960
Db 1672 ACTGAACAAGATTATACAACTATGCTGACCTTAAGCCCAACAAAGAGTACCCATT 1731
Qy 961 CTTTCTTTTGTATCAGACGAGGAGTCTAGGCGAGACTAGGTACTGGCATTTGGCGGTATC 1020
Db 1732 CTTTCTTTTGTATAGGACGAGGAGTCTAGGCGAGACTAGGTACTGGCATTTGGCGGTATC 1791
Qy 1021 ACAACCTCTACTAGTCTACTCAAACTATCTCAAGAAATAAATGTCACATGGAACAG 1080
Db 1792 ACAACCTCTACTAGTCTACTCAAACTATCTCAAGAAATAAATGTCACATGGAACAG 1851
Qy 1081 GTCACCTGACCTCCCTGGTCACTTTGCAAGATCAAACTTAACTCCCTAGCAGCAGTAGTCTTT 1140
Db 1852 GTGGCGACTCCCTGGTCACTTTGCAAGATCAAACTTAACTCCCTAGCAGCAGTAGTCTTT 1911
Qy 1141 CAAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTCTGTTTATTTTAA 1200

Db 1912 CAAAATCGAAGAGCTTTAGACTTTGCTAAACCGCTGAAAGAGGGGAAACCTGTTTATTTTAA 1971
Qy 1201 GGAGAAGAACGCTGTTTATTTATTTATGTAATCAATCCAGAAATGCTCACTGAGAAAGTTAAAGAA 1260
Db 1972 GGGAGAAGATGCTGTTTATTTATTTATGTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGAA 2031
Qy 1261 ATTGAGATCGAATACAAATGTAGACGAGAGGCTTCAAAACACCGAAGCTGGGGGCTC 1320
Db 2032 ATTGAGATCGAATACAAATGTAGACGAGAGGCTTCAAAACACCGAAGCTGGGGGCTC 2091
Qy 1321 CTAGCCCAATGGATGCCCTGGGTTCTCCCTTCTTAGGACCTCTAGGAGCTCTAATAATTG 1380
Db 2092 CTCAGCCAATGGATGGCTGGGTTCTCCCTTCTTAGGACCTCTAGGAGCTCTAATAATTG 2151
Qy 1381 TTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTAAAGTTTGTCTCTCCAGAAAT 1440
Db 2152 CTACTCTCTTTGGACCTCTGATCTTTAACTCTCTTTAACTCTCTTTAACTCTCTCTCCAGAAATC 2211
Qy 1441 GAAGCTCTAAAGCTTACAGATGGTCTTTACAAATGGAACCCCA 1481
Db 2212 GAAGCTCTAAAGCTTACAAATGGAGCCCAAGATGAGTCCAA 2252

RESULT 4
CR625046
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

2748 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DI044YK06 of Placenta Cot 25-normalized
of Homo sapiens (human).
CR625046
CR625046.1 GI:50505853
HTC; CNSLT_cDNA.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (Bases 1 to 2748)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue
2 (Bases 1 to 2748)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .2748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI044YK06"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 90.1%; Score 1333.8; DB 3; Length 2748;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 1 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTTCTTACCCCTTTGCTCTCACT 60
Db 783 ATGGCCCTCCCTTATCATATCTTTCTCTTTACTGTTCTTTTACCCCTTTTCACTCTCACT 842
Qy 61 GCACCCCTCCATGCTGCTGTACACCACTAGTCCCTTACCAAGAGTTTCTATGAAGA 120


```

843 GCACCCCTCCATGCGCTGTATGACCAAGTAGCTCCCTTTACCAAGAGTTTCTATGGAGA 902
121 AGCGGCTTCTGGAAATATTGATGCCCCCATATATAGAGTTTATCTAAGGGAACCTCC 180
903 ATGCAGCGTCCCGGAAATATTGATGCCCCCATGTATAGAGTCTTTCTAAGGGAACCCCC 962
181 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAAATCTCTCCACTCTTTTGCATG 240
963 ACCTTCACCTGCCACACCCATATGCCCCGCAACTGCTATAAATCTCTCCACTCTTTTGCATG 1022
241 CATGCAAACTACTCATTTATGGACAGGGAATAATTAATCTCTAGTTGTCTCTGAGGACTT 300
1023 CATGCAAACTACTCATTTATGGACAGGGAATAATTAATCTCTAGTTGTCTCTGAGGACTT 1082
301 GGAGCACTGTCTGTGGACTTACTTCAACCATACAGATATGCTCTGATGGGGTGAATTT 360
1083 GGAGTCACTGTCTGTGGACTTACTTCAACCAACTGGTATGCTCTGATGGGGTGGAGTT 1142
361 CRAAGGTCAGCAAGAGAAAACAAAGTAAAGGAAGCAATCTCCCAACTGACCCCGGGGACAT 420
1143 CAAGATCAGGCAAGAGAAAACATGTAAAGAGATTAATCTCCCAACTCACCAGGTTACAT 1202
421 AGCACCCCTAGCCCTACAAAGGACTAGTTTCTCTCAAAATACATGAAACCCCTCCGTACC 480
1203 GGCACCTCTAGCCCTACAAAGGACTAGATCTCTCAAAATACATGAAACCCCTCCGTACC 1262
481 CATATCTGCTGTGAGCCTATTATATACACCTCTACTCGGCTCCATGAGGTCTCAGCC 540
1263 CATATCTGCTGTGAGCCTATTATATACACCTCTACTCGGCTCCATGAGGTCTCAGCC 1322
541 CAATAACCTACTAATCTGTGGATGCTGCTCCCTCGACCTTCAAGCCATACATTTCAATC 600
1323 CAATAACCTACTAATCTGTGGATGCTGCTCCCTCGACCTTCAAGCCATATGTTCAATC 1382
601 CCTGTTCTCTGAACAATGGAACAACTTCAGACAGAAATAAACCACTTCCGTTTGTAGTA 660
1383 CCTGTTCTCTGAACAATGGAACAACTTCAGACAGAAATAAACCACTTCCGTTTGTAGTA 1442
661 GGAACCTCTTGTTCGAATCTGGAATAATCCCATACCTCAAACTCAGCTCTGTGTAATTT 720
1443 GGAACCTCTTGTTCGAATCTGGAATAATCCCATACCTCAAACTCAGCTCTGTGTAATTT 1502
721 AGCAATCTATAGACACAAACGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGA 780
1503 AGCAATCTATACACAAACGCTCCCAATGATCAGGTGGGTAAACCTCCACACGA 1562
781 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTG 840
1563 ATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGTACCTCAGCTATCATTTGTTG 1622
841 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTAGTGGCCCCCTATGACCATCTAC 900
1623 AATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTAGTGGCCCCCTATGACCATCTAC 1682
901 ACTGAAACAGATTTTATCAATCATGTCGTACCTAAGCCCCACACAAAGAGTACCCATT 960
1683 ACTGAAACAGATTTTATCAGTTATGTATATCTAAGCCCCGCAACAAAGAGTACCCATT 1742
961 CTTCTTTTGTATCAGACGAGGAGTGTAGGCACTAGTGTAGTCTGCGCATTCAGGATTC 1020
1743 CTTCTTTTGTATAGGACGAGGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGT 1802
1021 ACAACCTCTACTAGTTTCTACTACAAATATCTCAAGAAATAAATGTTGATGACATGGAACAG 1080
1803 ACAACCTCTACTAGTTTCTACTACAAATATCTCAAGAAATAAATGTTGATGACATGGAACAG 1862
1081 GTCACTGACTCCCTGGTCACTTCGACATCAATTAACCTCCCTAGCAGAGTGTAGTCTCT 1140
1863 GTCCGCACTCCCTGGTCACTTCGACATCAATTAACCTCCCTAGCAGAGTGTAGTCTCT 1922
1141 CAAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAAAGAGGGGAACTGTTTATTTTGA 1200
1923 CAAATCGAAGAGCTTTAGACTTGTCTAACCCGCTGAAGAGGGGAACTGTTTATTTTGA 1982

RESULT 5
CR617248 2749 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0D1022YJ18 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR617248.1 GI:50498055
VERSION HTC; CNSLT_CDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2749)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2749)
Genoscope.
Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES             Location/Qualifiers
     source           1..2749
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0D1022YJ18"
                     /tissue_type="Placenta Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      90.1%; Score 1333.8; DB 3; Length 2749;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY      1 ATGGCCCTCCCTTATCATATCTTTTCTTTTACTGTCTCTTACCCCTTTCGCTCTCACT 60
DB      777 ATGGCCCTCCCTTATCATATCTTTTCTTTTACTGTCTCTTTTACCTCTTCACTCTCACT 836
QY      61 GCACCCCTCCATGCTGCTGTACAAACAGTAGCTCCCTTACCAAGAGTTTCTATGAGA 120
DB      837 GCACCCCTCCATGCTGCTGTATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGAGA 896
```


Db	874	TTGCTCTGAGGACTTGAAGTCACTGCTGTTGGACTTACTTCAACCAACTGGTATGTC	815
Qy	345	TGATGGGGTGGAAATCAAGGTGAGCAAGAGAAACCAAGTAAGGAACCAATCTCCCA	404
Db	814	TGATGGGGTGGAGTTCAAGATCAGGCAAGAAACCAATGTAAGAGATGATTTCTCCCA	755
Qy	405	ACTGACCCGGGACATAGACACCCCTGACCCCTCAAAAGACTAGTTCTCTCAAAACTACA	464
Db	754	ACTACCCGGGTACATGGACCTCTAGCCCTCAAGACNCTAGATCTCTCAAACTACA	695
Qy	465	TGAACCCCTCCGACCATACCTCCCTGCTGAGCCCTATTATTAACAACCCCTACTCGGCT	524
Db	694	TGAACCCCTCCGACCATACCTCCCTGCTGAGCCCTATTATTAACAACCCCTACTCGGCT	635
Qy	525	CGATGAGTCTCAGCCCAACCCCTACTACTGTTGGATGTGCTCCCTGCTGACTTCAG	584
Db	634	CGATGAGTCTCAGCCCAACCCCTACTACTGTTGGATGTGCTCCCTGCTGACTTCAG	575
Qy	585	GCCATACATTTCAATCCCTGCTTCTGAAACCAATGAAACCACTTCAGACAGAAATAACAC	644
Db	574	GCCATATGTTTCAATCCCTGCTTCTGAAACCAATGAAACCACTTCAGACAGAAATAACAC	515
Qy	645	CACCTCCGTTTATAGTAGGACCTCTGTTTCCCAATCTGGAATTAACCCATACCTCAAACT	704
Db	514	CACCTCCGTTTATAGTAGGACCTCTGTTTCCCAATCTGGAATTAACCCATACCTCAAACT	455
Qy	705	CACCTGTGTAATTTAGCAATACATATAGACACACACAGCTCCCAATGCATCAGTGGGT	764
Db	454	CACCTGTGTAATTTAGCAATACATATAGACACACACAGCTCCCAATGCATCAGTGGGT	395
Qy	765	AACACCTCCACAGCAATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC	824
Db	394	AATCTCTCCACCAAAATAGTCTGCTTACCTCAGGAATATTTTGTCTGTGCTACCTC	335
Qy	825	AGCTATCATTTGTAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGC	884
Db	334	AGCTATCATTTGTAATGGCTCTTCAGAAATCTATGTGCTTCTCTCATTTCTTAGTGC	275
Qy	885	CCCTATGACATCTACCTGAAACAGATTTATACATCATGCTGCTACCTAAGCCCAAA	944
Db	274	CCCTATGACATCTACCTGAAACAGATTTATACATCATGCTGCTACCTAAGCCCAAA	215
Qy	945	CAAAAGTAGTACCAATCTCTCTTTGTTATCAGACAGGAGTGTAGGACAGACTAGGTAC	1004
Db	214	CAAAAGTAGTACCAATCTCTCTTTGTTATCAGACAGGAGTGTAGGACAGACTAGGTAC	155
Qy	1005	TGGCATTTGGCAGTATCAAAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	1064
Db	154	TGGCATTTGGCAGTATCAAAACCTCTACTCAGTTCTACTACAAACTATCTCAAGAACTAAA	95
Qy	1065	TGGTACATGGAACAGGTCTACTGCTCCCTGGTCACTTGCAGATCAACTTAACTCCCT	1124
Db	94	TGGGACATGGAACGGGTCTCCGACTCCCTGGTCACTTGCAGATCAACTTAACTCCCT	35
Qy	1125	AGCAGCAGTACTCTTCAAAATCGAAGCTTTA	1158
Db	34	AGCAGCAGTACTCTTCAAAATCGAAGCTTTA	1

RESULT 7
 BX347952
 LOCUS BX347952 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CSODI051Y13 5-PRIME, mRNA sequence.
 ACCESSION BX347952
 VERSION BX347952.1 GI:30375235
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 955)

AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?e=CS0BARE0232F07_AB02167_1&c=4215.r
FEATURES	Location/Qualifiers 1..955 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODI051Y13" /tissue_type="PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN	Query Match 54.9%; Score 813.6; DB 5; Length 955; Best Local Similarity 93.7%; Pred. No. 2e-229; Matches 890; Conservative 0; Mismatches 56; Indels 4; Gaps 4; Qy 226 GCCACTCTTTGGATGATGCAATATCTTATTTGGACAGGAAATGATTATCTTAGT 285 Db 1 GCCACTCTTTGGATGATGCAATATCTTATTTGGACAGGAAATGATTATCTTAGT 60 Qy 286 TGTCTCGAGGACTTGGAGCCACTGTCTGTGGACTTACTTCAACCCATACAGTATGTCT 345 Db 61 TGTCTCGAGGACTTGGAGCTTCTGTGTGGACTTACTTCAACCCATACAGTATGTCT 120 Qy 346 GATGGGGTGGAAATCAAGTTCAGGACAGAGAAAAAACAAGTAAGGAACAACTCCCAA 405 Db 121 GATGGGGTGGAGTTCAAGATCAGGCAAGAAAAACATGTAAAGAAAGTAATCTCCAA 180 Qy 406 CTGACCCGGGACATAGCACTTACCCCTAGCCCTACAAAGGACTAGTCTCTCAAACTACAT 465 Db 181 CTCACCCGGGTACATGGACCTCTAGCCCTACAAAGGACTAGTCTCTCAAACTACAT 240 Qy 466 GAAACCTCTCGTACCCATCTCGCCCTGGTGGAGCTATTTAATACCCCTCACTCGGCTC 525 Db 241 GAAACCTCTCGTACCCATCTCGCCCTGGTGGAGCTATTTAATACCCCTCACTCGGCTC 300 Qy 526 CATGAGTCTCAGCCCCAAAACCTTAACTGTTGATGTGCTTCCCTTGCATCTCAGG 585 Db 301 CATGAGTCTCAGCCCCAAAACCTTAACTGTTGATGTGCTTCCCTTGCATCTCAGG 360 Qy 586 CCATACATTTCAATCTCTGTTCTGACATGCAACCACTTTCAGCAGAGAAATAACACC 645 Db 361 CCATATGTTTCAATCTCTGTTCTGACATGCAACCACTTTCAGCAGAGAAATAACACC 420 Qy 646 ACTTCCGTTTCTAGTAGGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTC 705 Db 421 ACTTCCGTTTCTAGTAGGACCTCTGTTTCCAAATCTGGAATAACCCATACCTCAAACTC 480 Qy 706 ACCTGTGTAATAATTTAGCAATATCTATAGACAAACCAAGCTCCCAATGCATCAGTGGGA 765 Db 481 ACCTGTGTAATAATTTAGCAATATCTATAGACAAACCAAGCTCCCAATGCATCAGTGGGA 540 Qy 766 ACACCTCCACAGCAATAGTCTGCTACCTCTAGGATATTTTGTCTGTGGTACCTCA 825 Db 541 ACTCCT-CCAAACAAATAGTCTGCTACCTCTAGGATATTTTGTCTGTGGTACCTCA 599

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 872)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAD0102C11_AD00948_1&c=4215.r

FEATURES
source Location/Qualifiers
1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized." DB 5; Length 872;

ORIGIN
Query Match 50.5%; Score 747.2; DB 5; Length 872;
Best Local Similarity 92.3%; Pred. No. 9.7e-210;
Matches 808; Conservative 0; Mismatches 64; Indels 3; Gaps 2;
278 ATCTAGTGTCTGGAGGACTTGGAGCCACTGCTCTGGAGCTACTTACCCCATACCA 337
872 ATCTAGTGTCTGGAGGACTTGGAGCCACTGCTCTGGAGCTACTTACCCCATACCT 813
338 GTATGCTGATGGGGTGGATTCAAGTTCAGGCAAGAGAAAACAGTAAGAAACAA 397
812 GGTATGCTATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAACAGTAAGAA- 754
398 TCTCCCAACTGACCCGGGACATAGCACCCCTAGCCCTCAAAAGGACTAGTCTCTCAA 457
753 TCTCCCAACTGACCCGGGACATAGCACCCCTAGCCCTCAAAAGGACTAGTCTCTCAN 694
458 AACTACATGAACCCCTCGTACCCACTACTCGCTGGTGAGCCTATTAAATACCACTCA 517
693 AACTACATGAACCCCTCGTACCCACTACTCGCTGGTGAGCCTATTAAATACCACTCA 634
518 CTCGGCTCCATGAGGTCTACCCCAAAACCTACTACTGCTGAGTGTGCTCCCTGTC 577
633 CTGGGCTCCATGAGGTCTCGGCCCAAAACCTACTACTGCTGAGTGTGCTCCCTGTA 574
578 ACTTCAGGCCATACATTTCAATCTCTGTTCTGAAACATGAACAACTTCAGACAGAAA 637
573 ACTTCAGGCCATATGTTTCAATCTCTGTTCTGAAACATGAACAACTTCAGACAGAAA 514
638 TAAACACCACTTCGGTTTATAGGACCTCTGTTTCCAAATCTGGAATAACCCATACCT 697
513 TAAACACCACTTCGGTTTATAGGACCTCTGTTTCCAAATCTGGAATAACCCATACCT 454
698 CAAACCTCACTGTGTAATAATTTAGCAATATATAGACACAAACAGCTCCCAATGCATCA 757
453 CAAACCTCACTGTGTAATAATTTAGCAATATATAGACACAAACAGCTCCCAATGCATCA 394
758 GGTGGGTAAACCTCCACAGCATAGTCTGCTACCTCAGGAATATTTTCTGCTGTG 817

Db 393 GGTGGGTAACTCCTCCACACAAAATAGTCTGCTACCTACCCCTCAGGAATATTTTGTCTGTG 334
Qy 818 GTACCTCAGCCTATCATTTGTAATGGCTCTTTCAGAACTATGTGCTTCTCTCATTTCT 877
Db 333 GTACCTCAGCCTATCGTTGTTGAATGGCTCTTTCAGAACTATGTGCTTCTCTCATTTCT 274
Qy 878 TAGTGGCCCTATGACCATCTACACTGTAACAAGATTTATACAGTATTATGTCATATCTAAGC 937
Db 273 TAGTGGCCCTATGACCATCTACACTGTAACAAGATTTATACAGTATTATGTCATATCTAAGC 214
Qy 938 CCCACAAAAGAGTACCATCTTCTTCTTGTATCATCAGACGAGGTGCTAGGCGAGAC 997
Db 213 CCGGCAACAAAAGAGTACCATCTTCTTCTTGTATAGAGGAGGAGTCTAGGTGCGAC 154
Qy 998 TAGGTACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAG 1057
Db 153 TAGGTACTGGCATTTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTATCTCAAG 94
Qy 1058 AATAAATGGTGAATGGAACAGGTCTACTGCTCTGCTGCTGCTGCAAGATCAACTTA 1117
Db 93 AACTAAATGGGACATGGAACGGTCTCGGACTCC--TGGTCACTGCAAGATCAACTTA 36
Qy 1118 ACTCCCTAGCAGCAGTAGTCTTCAAAATCGAGA 1152
Db 35 ACTCCCTAGCAGCAGTAGTCTTCAAAATCGAGA 1

RESULT 10
BX408752 903 bp mRNA linear EST 01-MAY-2004
BX408752 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120
5-PRIME, mRNA sequence.
ACCESSION BX408752
VERSION BX408752.2 GI:46931100
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 903)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30648825.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAA0082H03_CS00697_2&c=4215.r

FEATURES
source Location/Qualifiers
1..903
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013Y120"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized." DB 5; Length 903;

ORIGIN
Query Match 50.0%; Score 741; DB 5; Length 903;

```
Best Local Similarity 93.0%; Pred. No. 6.8e-208;
Matches 796; Conservative 0; Mismatches 58; Indels 2; Gaps 2;

QY 4 GCCCTCCCTATACATACATCTTTCTTTACTGTCTCTTACCCCTTTTCGCTCTACTGCA 63
Db 50 GGCCCTCCCTTANATATTTTCTCTTACTGTCTTTTACCCCTTTTCACTCTCACTGCA 109
QY 64 CCCCTCCATGCTGTGTACACCAAGTAGCTCCCTTACCAAGAGTTTCTATGAAGACG 123
Db 110 CCCCTCCATGCGGTGTATGACCAAGTAGCTCCCTTACCAAGAGTTTCTATGGAATG 169
QY 124 CGCTTCTCGGAATATGATGCCCATCATATAGGAGTTTATCTAAGGAACTCCACC 183
Db 170 GAGCGTCCGGAATATGATGCCCATCATATAGGAGTTTCTAAGGAACTCCACC 229
QY 184 TTCCTGCCCCACCCATATGCCCCGCAACTGCTATAACTCTGCACTCTTTGCAATGAT 243
Db 230 TTCCTGCCCCACCCATATGCCCCGCAACTGCTATAACTCTGCACTCTTTGCAATGAT 289
QY 244 GCAATACTCATATTTGGACAGGGAATGATTAATCTAGTTGCTCTGGAGGACTTGA 303
Db 290 GCAATACTCATATTTGGACAGGGAATGATTAATCTAGTTGCTCTGGAGGACTTGA 349
QY 304 GGCCTGCTGTGTGGACTTTACTTCCCATACCAAGCAATCTCTGATGGGGTGGAAITCAA 363
Db 350 GTCCTGCTGTGTGGACTTTACTTCCCAAACTGGTAATGCTGATGGGGTGGAGTTCAA 409
QY 364 GTCAGGCAAGAAAGAAACAAAGAAAGCAATCTCTGATGGGGTGGAAITCAA 423
Db 410 GATCAGGCAAGAAAGAAACATGTAAGAAAGTAATCTCCCAACTCACCCTGGGTACATGGC 469
QY 424 ACCCTAGCCCTACAAAGGAGTGTCTCTCAAACTACATAAGAAACCTCCGTACCCAT 483
Db 470 ACTCTAGCCCTTACAAAGGAGTGTCTCTCAAACTACATAAGAAACCTCCGTACCCAT 529
QY 484 ACTCGCTGCTGAGCCTATTTAATACCAACCCTCACTCGGCTCCATGAGTCTTCAGCCAA 543
Db 530 ACTCGCTGCTGAGCCTATTTAATACCAACCCTCACTCGGCTCCATGAGTCTTCAGCCAA 589
QY 544 ACCCTACTAACTGTGGATGTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603
Db 590 AACCTTACTAACTGTGGATGTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
QY 604 GTTCTGAAACATGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTAGTAGA 663
Db 650 GTACCTGAAACATGGAACAACTTTCAGCAGAGAAATAAACACCACTTCCGTTTAGTAGA 709
QY 664 CCTCTTGTGTTTCCAACTGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 723
Db 710 CCTCTTGTGTTTCCAACTGGAATAAACCCATACCTCAAACTCACCTGTGTAAATTTAGC 769
QY 724 AATACATAGACAAACAGCTCCCAATGCATCAGTGGGTAAACACCTCCACCAAGATA 783
Db 770 AATACATAGACAAACAGCTCCCAATGCATCAGTGGGTAAACACCTCCACCAAGATA 829
QY 784 GTCTGCTACCTCCAGGAATATTTTGTGCTGTGGTACCTCAGCTCATATGTTTGAAT 843
Db 830 GTCTGCTACCTCCAGG-ATAATTTTGTGCTG-GGTACCTCAGCTTATTCGTTTGGAT 887
QY 844 GGCTCTTCAGAAATCTA 859
Db 888 GGCTCTTCAGAAATCTA 903
```

```
RESULT 11
BX408733
LOCUS
DEFINITION
BX408733 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YJ24
5-PRIME, mRNA sequence.
ACCESSION
BX408733
VERSION
BX408733.1 GI:30635957
KEYWORDS
EST.
Homo sapiens (human)
SOURCE
```

```
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 883)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAA0082E02_CS00686_l&c=4215.r
.
FEATURES
source
1..883
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 49.8%; Score 737.2; DB 5; Length 883;
Best Local Similarity 93.2%; Pred. No. 9e-207;
Matches 825; Conservative 0; Mismatches 54; Indels 6; Gaps 5;
QY 29 TTACTGTCTCTTACCCCTTTTCGCTCTCACTGCACCCCTCCATGCTGTGTACAAACA 88
Db 1 TTACTGTCTCTTACCCCTTTTCGCTCTCACTGCACCCCTCCATGCTGTGTACAAACA 58
QY 89 GTAGCTCCCTTACCAAGAGTTTCTATGAAGAACGCGGCTTCTGGAAATATGATGCC 148
Db 59 GTAGCTCCCTTACCAAGAGTTTCTATGAAGAACGCGGCTTCTGGAAATATGATGCC 118
QY 149 CATCATATAGAGTTTATCTAAGGAACTCCACTTCACCTGCCACACCCATATGCC 208
Db 119 CATCGTATAGAGTCTTTCTAAGGAAACCCCTTCACTGCGCCACACCCATATGCC 178
QY 209 GCAACTGTATACTTGCCTCTTTCATGATGCAATGCAATACCTATTATTGGACAGGA 268
Db 179 GCAACTGTATACTTGCCTCTTTCATGATGCAATGCAATACCTATTATTGGACAGGA 238
QY 269 AAATGATTAACTCTAGTTGTCTCTGAGGACTTGGAGCACTGTCTGTGGACTTACTTCA 328
Db 239 AAATGATTAACTCTAGTTGTCTCTGAGGACTTGGAGCACTGTCTGTGGACTTACTTCA 298
QY 329 CCCATACAGATGTCTGATGGGGTGGAAATTCAGGTTCAGGCAAGAGAAACAAGTAA 388
Db 299 CCCAACTGGTATGTCTGATGGGGTGGAGTTCAAGATTCAGGCAAGAGAAACAAGTAA 358
QY 389 AGGAAGCAATCTCCCACTGACCGGGGACATAGCACCCCTAGCCCTTACAAAGGACTAG 448
Db 359 AAGAAGTAACTCTCCCACTGACCGGGGACATAGCACCCCTTAGCCCTTACAAAGGACTAG 418
QY 449 TTCTCTCAAACTACATGAACCCCTCCGTACCCATACCTGCTGCTGAGGCTATTTAATA 508
Db 419 ATCTCTCAAACTACATGAACCCCTCCGTACCCATACCTGCTGCTGAGGCTATTTAATA 478
QY 509 CCACCTCACTCGGCTCCATGAGTCTCAGGCCCAACCCCTACTAACTGTTGGATGCC 568
Db 509 CCACCTCACTCGGCTCCATGAGTCTCAGGCCCAACCCCTACTAACTGTTGGATGCC 568
```



```

Db      479  CCACCCTCACTGGGCTCCATGAGGTCTCGCGCCAAAACCTACTAACTGTTGATATGCC 538
Qy      569  TCCCCCTGCACTTCAGGCCATACATTTCAATCCCTGTTCTGACCAATGGAACAACCTTCA 628
Db      539  TCCCCCTGAACCTTCAGGCCATATGTTTCAATCCCTGCTACTGAAACCAATGGAACAACCTTCA 598
Qy      629  GCACAGAAATAACACACACCTTCGGTTTATAGTAGGACTCTTGTGTTCCAAATCTGGAATAA 688
Db      599  GCACAGAAATAACACACCTTCGGTTTATAGTAGGACTCTTGTGTTCCAAATCTGGAATAA 658
Qy      689  CCCATACCTCAAAACCTCACTCGTGTGTAATAATTTAGCAATACTATAGACACAACAGCTCCC 748
Db      659  CCCATACCTCAAAACCTCACTCGTGTGTAATAATTTAGCAATACTATAGACACAACCAACTNCC 718
Qy      749  AATGCAATCA-GGTGGGTAAACACCTCCACACGAAATAGTCTGCTACCTCAGGAATATT 807
Db      719  AATGCAATCAAGGTGGGTAAACCTCCCTCCACAAATAAGTCTGCTACCTCA-GAATATT 777
Qy      808  TTTGCTGTGTGCTACCTCAGCTCATATGTTTGAATGGCTC-TTCAGAATCTATGTGCTT 866
Db      778  TTTGCTGTGTGCTACCTCAGCTCATATGTTTGAATGGCTCTTTCAGAATCTATGTGCTT 837
Qy      867  CCTCTCATCTTATAGTGGCCCC-TATGACCACTATACACTGAACAAG 910
Db      838  CCTCTCATCTTATAGTGGCCCCCTATGACATTTACACTGAACAAG 882

RESULT 12
BX347619/c      877 bp      mRNA      linear      EST 23-APR-2004
LOCUS      BX347619 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      clone CSOD1022XJ18 3-PRIME, mRNA sequence.

ACCESSION      BX347619
VERSION      BX347619.1
KEYWORDS      GI:30379218
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 877)
AUTHORS      Li, W.B., Gruber, C., Jesses, J., and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAD0092B06_AD00811_1&c=4215.r

```

FEATURES

```

source      Location/Qualifiers
1..877
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1022XJ18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match      49.2%; Score 728.4; DB 5; Length 877;
Best Local Similarity 93.2%; Pred. No. 3.7e-204;
Matches 803; Conservative 0; Mismatches 55; Indels 4; Gaps 4;

```

```

Qy      503  TTAATACCAACCCCTCACTCGGCTCCATGAGGTCTC-AGCCCAAAACCCCTACTAATCTGTTGG 561
Db      877  TAAACCAACCCCTCACTCGGCTCCATGAGTNCCTCGGGCCAAAACCCCTACTAATCTGTTGG 818
Qy      562  ATGTGCTCCCTCGCTGCACTTCAGGCCATACATTTCAATCCCTGTTCTCGAACAATGGAAC 621
Db      817  ATATGCTCCNCCTGAACCTTCAGGGCATATGTTTCAATCCCTGTACTCGAACAATGGAAC 758
Qy      622  AACTTCAGACACAGA-AATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTGTTCCAAATCT 680
Db      757  ACTTTCAGACAGANATATAACACCACTNCCGTTTTAGTAGGACCTCTTGTGTTCCAAATCT 698
Qy      681  GGAATAAACCCTACCTCAAACTCCTGCTGTGTAAATTTAGCAATACTATAGACACAAC 740
Db      697  GGAATAAACCCTACCTCAAACTCCTGCTGTGT-AAAATTTAGCAATACTATAGACACAAC 639
Qy      741  CAGCTCCCAATGATCAGGTGGGTAAACCTCCACACGAATAAGTCTGCTACCTCCCTCAGG 800
Db      638  CAACTCNCATGATCAGGTGGGTAACTCTCTCCACACAATAAGTCTGCTACCTCCCTCAGG 579
Qy      801  AATATTTTGTGCTGTGGTACCTCAGCCTATCATTTGTTGAATGGCTCTTTCAGAATCTAT 860
Db      578  AATATTTTGTGCTGTGGTACCTCAGCCTATCGTTGTTGAATGGCTCTTTCAGAATCTAT 519
Qy      861  GTGCTTCTCTCATCTTATAGTGGCCCCCTATGACCATCTACACTGAAACAAGATTATACAA 920
Db      518  GTGCTTCTCTCATCTTATAGTGGCCCCCTATGACCATCTACACTGAAACAAGATTATACAG 459
Qy      921  TCAATGCTGTAACCTTAAGCCCAACAAAGAGTACCATCTTCTTCTTGTGTTATCAGGC 980
Db      458  TTAATGTCATATCTAAGCCCAACAAAGAGTACCATCTTCTTCTTGTGTTATAGGAGC 399
Qy      981  AGGAGTGCTAGGCAGACTAGGTACTGTCATTTGGCAGTATCACAACTCTACTCAGTTCTA 1040
Db      398  AGGAGTGCTAGGTGCTAGGTACTGTCATTTGGCAGTATCACAACTCTACTCAGTTCTA 339
Qy      1041  CTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTGGTGCAC 1100
Db      338  CTACAAACTATCTCAAGAAATAAATGGGACATGGAACGGTCCGCGCTCCCTGGTGCAC 279
Qy      1101  CTTGCAAGATCAACTTAACCTCCCTAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGA 1160
Db      278  CTTGCAAGATCAACTTAACCTCCCTAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGA 219
Qy      1161  CTTGCTAACCGCAAAAGAGGGGAACTGTTATTTTATAGGAGAAACGCTCTATTATA 1220
Db      218  CTTGCTAACCGCTGAAAGAGGGGAACTGTTATTTTATAGGAGAAAGATGCTGTTATTA 159
Qy      1221  TGTTAATCAATCAGAAATTTGCTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATG 1280
Db      158  TGTTAATCAATCGGAATCTGCTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAACG 99
Qy      1281  TAGAGCAGAGGAGCTTCAAAACACCGAACCGCTGGGGCTCTCTCAGCAATGGATGCCCTG 1340
Db      98  TAGAGCAGAGGAGCTTCAAAACACCTGGACCTCGGGCTCTCTCAGCAATGGATGCCCTG 40
Qy      1341  GGTTCCTCCCTCTTAGGACCT 1362
Db      39  GATTCCTCCCTCTTAGGACCT 18

```

RESULT 13

```

BX388766/c      836 bp      mRNA      linear      EST 29-APR-2004
LOCUS      BX388766 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      clone CSOD104YK06 3-PRIME, mRNA sequence.

ACCESSION      BX388766
VERSION      BX388766.2
KEYWORDS      GI:46877917
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 836)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30462383.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AU0122B04_U01082_l&c=4215.r.

FEATURES
source

1..836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI044YK06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 48.2%; Score 714; DB 5; Length 836;
Best Local Similarity 92.6%; Pred. No. 6.8e-200;
Matches 772; Conservative 0; Mismatches 60; Indels 2; Gaps 2;

QY 565 TGCCCTCCCTGACCTTCAGGCGCATACATTTCAATCCCTGTTCCCTGAACAATGGAAACAC 624
DB 835 TGCCCTCCCTGACCTTCAGGCGCAATGTTCAATCCCTGTTCCCTGAACAATGGAAACAC 776
QY 625 TTCCAGCACAGAAA-TAAACACCACTTCGTTTTAGTAGGACCTCTGTTTCCAACTCGGA 683
DB 775 TTCCAGCCAGAAATTACACACCACTCCCGTTTAAAGTAGGACCTCTGTTTCCAACTCGGA 716
QY 684 AATAACCCATACCTCAAACTCAGCTGTGTAAATTTAGCATATCTATAGACACACACAG 743
DB 715 AATAACCCATTCCTCAAACTCAGCTGTGTAAATTTAGCAATTCCTACATACACACCCAA 656
QY 744 CTCCCAATGCATCAGTGGGTAAACACCTCCACACAGATAGTCTGCCTACCCCTCAGGAAT 803
DB 655 CTCCCAATGCATCAGTGGGTAACTCTCTCCACACAAATAGTCTGCCTACCCCTCAGGAAT 596
QY 804 ATTTTGTGTGTGTGTTACCTCAGCCCTATCATTTGTTTGAATGCTCTTTCAGATCTATGTG 863
DB 595 ATTTTGTGTGTGTGTTACCTCAGCCCTATCATTTGTTTGAATGCTCTTTCAGATCTATGTG 536
QY 864 CTTCCTCTCATTTCTAGTGGCCCTATGACCATCTACACTGACACAGATTTTACAACTCA 923
DB 535 CTTCCTCTCATTTCTAGTGGCCCTATGACCATCTACACTGACACAGATTTTACAGTTA 476
QY 924 TGTGCTGACTAGCCCCACAAACAAAGAGTAGCCATCTCTCTTTGTTTATCAGACGAGG 983
DB 475 TGTGCTGACTAGCCCCACAAACAAAGAGTAGCCATCTCTCTTTGTTTATGAGGAGGAGG 416
QY 984 AGTGCTAGGACACTAGGTACTGGCAATGGCAGTATCAACAACCTCTACTCAGTTCTACTTA 1043
DB 415 AGTGCTAGGTGACACTAGGTACTGGCAATGGCAGTATCAACAACCTCTACTCAGTTCTACTTA 356
QY 1044 CAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTACTCCTCGTCACTTT 1103
DB 355 CAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTACTCCTCGTCACTTT 296
QY 1104 GCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTT 1163

Db 295 GCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTT 236
QY 1164 GCTAACCGCCCAAAGAGGGGAACTGTTTATTTTAGGAGAGAAAGCGCTGTTATTATGT 1223
Db 235 GCTAACCGCTGAAAGAGGGGAACTGTTTATTTTAGGAGAGAAAGCGCTGTTATTATGT 176
QY 1224 TAATCAATCCAGAAATTTGCTACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACACGTAG 1283
Db 175 TAATCAATCCGAAATGCTACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACACGTAG 116
QY 1284 AGCAGAGAGCTTCAAAACACCGAACGCTGGGGCTCTCTAGCCAAATGAGTGCCTGGGT 1343
Db 115 AGCAGAGAGCTTCAAAACACCTGACCTGGGGCTCTCTAGCCAAATGAGTGCCTGGAT 56
QY 1344 TCTCCCTCTTCTAGGACCTCTAGCAGCTCTAATATTTGTTACTCTCTTTGGACC 1397
Db 55 TCTCCCTCTTCTAGGACCTCTAGCAGCTCTAATA-TGCTACTCCACTTTGGGCCC 3

RESULT 14
EX408734
LOCUS

DEFINITION BX408734 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012J24
5-PRIME, mRNA sequence.
ACCESSION BX408734
VERSION BX408734.1 GI:30635959
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 921)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAA0082B02_CS00686_2&c=4215.r

FEATURES
source

Location/Qualifiers
1..921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012J24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 47.9%; Score 709.8; DB 5; Length 921;
Best Local Similarity 88.6%; Pred. No. 1.2e-198;
Matches 815; Conservative 0; Mismatches 98; Indels 7; Gaps 4;

QY 29 TTACTGTTCTCTTACCCCTTTTCCGCTCTCACTGCACCCCTCCATGCTGCTGTACAAACA 88
Db 1 TTCTGTTCTTTTACCCCTCTTTTCACTCTCACTGCACCCCTCCATGCTGCTGTATGACCA 60
QY 89 GTAGCTCCCTTTACCAAGAGTTTCTATGAAGAACGCGGCTTCTCTGGAAATATTGTAGCCC 148

Db 61 GTAGTCCCTTACCAANAGTTTCTATGGAGATGAGCGTCCCG- AATATTGATGCC 119
 Qy 149 CATCATATAGGAGTTTATCTAAGGAAATCTCCACCTTCTAGTCCCAACCCATATGCC 208
 Db 120 CATCGTATAGGAGTCTTCTAAGGAAACCCCACTTCTAGTCCCAACCCATATGCC 179
 Qy 209 GCAACTGTATATCTGACCTCTTGTGATGATGCAATATCTATTTAGGAGGGA 268
 Db 180 GCAACTGTATATCTGACCTCTTGTGATGATGCAATATCTATTTAGGAGGGA 239
 Qy 269 AATAGTATATCTGACCTCTTGTGATGATGCAATATCTATTTAGGAGGGA 328
 Db 240 AATAGTATATCTGACCTCTTGTGATGATGCAATATCTATTTAGGAGGGA 299
 Qy 329 CCATACCATATGTTCTGATGGGGTGGAAATCAAGTCAAGCAAGAAACAGTAA 388
 Db 300 CCCAAACTGGTATGTTCTGATGGGGTGGAGTTCAAGATCAGGCAAGAAACAGTAA 359
 Qy 389 AGAAGCAATCTCCCACTGACCCGGGAGATAGACCCCTAGCCCTCAAAAGGACTAG 448
 Db 360 AGAAGTAAATCTCCCACTGACCCGGGAGATAGACCCCTAGCCCTCAAAAGGACTAG 419
 Qy 449 TTCTCTCAAACTACATGAAACCTCGTACCCATATCTGCGTGTGAGCCTATTTAATA 508
 Db 420 ATCTCTCAAACTACATGAAACCTCGTACCCATATCTGCGTGTGAGCCTATTTAATA 479
 Qy 509 CCACCTCTACTCGGCTCCATGAGGTCTGAGCCCAAAACCTCTACTTGTGATGTGCC 568
 Db 480 CCACCTCTACTCGGCTCCATGAGGTCTGAGCCCAAAACCTCTACTTGTGATGTGCC 539
 Qy 569 TCCCTCTGACTTCAGGCAATATCTCAATCCCTGTTCTGAGCAATGGAACCACTTCA 628
 Db 540 TCCCTCTGACTTCAGGCAATATGTTCAATCCCTGTTCTGAGCAATGGAACCACTTCA 599
 Qy 629 GCACAGAAATAAACACCACTTCCGTTTGTAGGACCTCTTGTTCCTGGAATAAA 688
 Db 600 GCACAGAAATAAACACCACTTCCGTTTGTAGGACCTCTTGTTCCTGGAATAAA 659
 Qy 689 CCCATACCTCAAACTCACTGTTGTAATTTAGCAATATCTAGACCAACCACTTCC 748
 Db 660 CCCATACCTCAAACTCACTGTTGTAATTTAGCAATATCTAGACCAACCACTTCC 718
 Qy 749 AATGATATAGGTTAGGTTAACTTCCCTCCCAAAATAGTCTGCTCAGGATATTTT 808
 Db 719 AATGATATAGGTTAGGTTAACTTCCCTCCCAAAATAGTCTGCTCAGGATATTTT 778
 Qy 809 TTGCTCTGTTAGTCTGAGCTATCATTTGTTGATGGCTCTTCAAGATCTATGCTTCC 868
 Db 779 TTGCTCTGTTAGTCTGAGCTATCATTTGTTGATGGCTCTTCAAGATCTATGCTTCC 837
 Qy 869 TCTCATTTAGTCTGAGCTATGACCTATGACCTGAAAGATTTATCAATCATGTCG 928
 Db 838 CTTATCTAGGCTGAGCTATGACCTATGACCTGAAAGATTTATCAATCATGTCG 893
 Qy 929 TACCTAAGCCCCCAACAAA 948
 Db 894 ATACTAAGCCCCCAACAAA 913

RESULT 15
 BX347111/c 844 bp mRNA linear EST 08-APR-2004
 LOCUS BX347111 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1022YE21 3-PRIME, mRNA sequence.
 ACCESSION BX347111
 VERSION BX347111.1 GI:30375145
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 844)

AUTHORS
 TITLE
 JOURNAL
 COMMENT

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7108.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?&s=CS0AU0102B08_U0922_1&c=7108.r.

Location/Qualifiers

source

1..844

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1022YE21"

/issue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 47.7%; Score 705.8; DB 5; Length 844;

Best Local Similarity 92.7%; Pred. No. 1.8e-197;

Matches 762; Conservative 0; Mismatches 58; Indels 2; Gaps 2;

Qy 502 TTTAATACCACTCTACTCGGCTCCATGAGGTCTGAGCCCAAAA-CCCTACTAATCTGTTG 560
 Db 837 TTAATATCCACCTCTCCCTGGTTTCCATTTAGGCTCGGCCCAAAACCCCTACTACTG 778
 Qy 561 GATGTGCTT-CCCTCTGACATTCAGGCCATACATTTCAATCCCTGTTCTGGAACAATGGA 619
 Db 777 GATATGCTTCCCTCCCTGNAATTCAGGCCATATGTTTCAATCCCTGTTACTGGAACAATGGA 718
 Qy 620 ACAATCTGAGCAGAGAAATAAACACCACTCCGTTTGTAGGACCTCTGTTTCCATC 679
 Db 717 CAACTTTGAGCAGAGAAATAAACCCCACTTCCGTTTGTAGTGGACCTCTTGTTCCTCAATC 658
 Qy 680 TGGAAATAACCCATACCTCAAACTCAGCTGTTGTAATTTAGCAATATCTATAGACAA 739
 Db 657 TGGAAATAACCCATACCTCAAACTCAGCTGTTGTAATTTAGCAATATCTATAGACAA 598
 Qy 740 CCAGCTCCCAATGCATCAGGTGGGTAAACCTCCACACGAAATAGTCTGCTACCTCAG 799
 Db 597 CCAGCTCCCAATGCATCAGGTGGGTAACTCTCCACACGAAATAGTCTGCTACCTCAG 538
 Qy 800 GAATATTTTGTGTGTGTAACCTGAGCTATCATTTGTTGATGGCTCTTCAAGATCTA 859
 Db 537 GAATATTTTGTGTGTGTAACCTGAGCTATCATTTGTTGATGGCTCTTCAAGATCTA 478
 Qy 860 TGTGCTTCTCTCATTTCTAGTGGCTTATGAGCTTATGAGCTTATGAGCTTATGAG 919
 Db 477 TGTGCTTCTCTCATTTCTAGTGGCTTATGAGCTTATGAGCTTATGAGCTTATGAG 418
 Qy 920 ATCATGTGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAG 979
 Db 417 GTTATGTATCATTTATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAGCTTATGAG 358
 Qy 980 CAGGAGTGTCTAGGAGCTAGGTACTGGCAATGGCAATGGCAATGGCAATGGCAATGGCA 1039
 Db 357 CAGGAGTGTCTAGGAGCTAGGTACTGGCAATGGCAATGGCAATGGCAATGGCAATGGCA 298
 Qy 1040 ACTACAACTATCTCAAGAAATAAATGGTGCATGGAACAGGCTCAGCTCCCTGCTCA 1099
 Db 297 ACTACAACTATCTCAAGAAATAAATGGTGCATGGAACAGGCTCAGCTCCCTGCTCA 238

```
QY 1100 CTTTGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAG 1159
|||
Db 237 CTTTGAAGATCACTTAACCTCCCTAGCAGCAGTAGTCTCTCAAAATCGAAGAGCTTTAG 178
|||
QY 1160 ACTTCTTAACCCCAAAAGAGGGGAACTGTTTATTTTATTTAGGAGAAAGAGCTGTTATT 1219
|||
Db 177 ACTTCTTAACCCCTGAAGAGGGGAACTGTTTATTTTATTTAGGAGAAAGAGCTGTTATT 118
|||
QY 1220 ATGTTAAATCAATCCAGAAATGTCACAGAAAGTTAAAGAAATTCGAGATCGAATACAAAT 1279
|||
Db 117 ATGTTAAATCAATCCGAAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAAC 58
|||
QY 1280 GTAGAGCAGAGAGCTTCAAAACCGAAGCTGCGGGCTCC 1321
|||
Db 57 GTAGAGCAGAGAGCTTCAAAACACTGGACCTGGGGGCGGCC 16
|||

RESULT 16
BX367907
LOCUS
DEFINITION
BX367907 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION
BX367907
VERSION
BX367907.1 GI:30459586
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 828)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AT022ZG01_T02023_1&c=4215.r.
Location/Qualifiers
1..828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match 47.5%; Score 703.2; DB 5; Length 828;
Best Local Similarity 93.8%; Pred. No. 1.1e-196;
Matches 754; Conservative 0; Mismatches 48; Indels 2; Gaps 2;

ORIGIN
QY 110 TTCTATGAAGACCGCGCTCTGGAATATTGATGCCCATCATATAGAGCTTATCTA 169
|||
Db 9 TTTCTTGGAGATGCGAGCGTCCGG--AATAATTGATGCCCATCGTAGAGTCTTTCTA 67
|||
QY 170 AGGGAACCTCCACTTCACTGCCACACCATATGCCCGCAACTGCTATACTCTGCCA 229
|||
Db 68 AGGGAACCCACCTTCACTTGCCACACCATATGCCCGCAACTGCTATCACTCTGCCA 127
|||
QY 230 CTCTTTGCATGTCGAAATACTCATTTATGGACAGGGAATGATTAATCTTAGTTGTC 289
|||
```

```
Db 128 CTCTTTGCATGTCGAAATACTCATTTATTGGACAGGAAAAATGATTAACTTAGTTGTC 187
|||
QY 290 CTGGAGACTTTGGAGCCCACTGCTGTTGGACTTACTTCAACCATACCAGTATGCTGATG 349
|||
Db 188 CTGGAGACTTTGGAGTCACTGCTGTTGGACTTACTTCAACCATACCAGTATGCTGATG 247
|||
QY 350 GGGGTGGAATTTCAAGGTTCAGGCAAGAGAAAAACAAGTAAAGGAAGCAATCTCCCAACTGA 409
|||
Db 248 GGGGTGGAATTTCAAGTTCAGGCAAGAGAAAAACAAGTAAAGGAAGTAAATCTCCCAACTCA 307
|||
QY 410 CCCGGGACATAGACACCCCTAGCCCTTCAAGAGCTAGTCTCTCAAAACTACATGAAA 469
|||
Db 308 CCCGGGTACATGCGCACCTCTAGCCCTTCAAGAGCTAGATCTCTCAAAACTACATGAAA 367
|||
QY 470 CCTTCGTTACCCATCTCGCTGCTGAGCCCTATTTAATACCCCTCACTCGGCTCCATG 529
|||
Db 368 CCTTCGTTACCCATCTCGCTGCTGAGCCCTATTTAATACCCCTCACTCGGCTCCATG 427
|||
QY 530 AGGTCTCAGCCCAAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCTCACTTCAGGCCAT 589
|||
Db 428 AGGTCTCGGCCCAAAACCCCTACTAACTGTTGGATGTCCTCCCTCGCTCACTTCAGGCCAT 487
|||
QY 590 ACATTTCAATCCCTGTTCTCGAACAATGGAACAACATTCAGCAGAGAAATAAACACCACTT 649
|||
Db 488 ATGTTTCAATCCCTGTTCTGGAACAATGGAACAACATTCAGCAGAGAAATAAACACCACTT 547
|||
QY 650 CCGTTTGTAGTAGGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCACCT 709
|||
Db 548 CCGTTTGTAGTAGGACCTCTGTTTCCAACTCTGGAATAAACCCATACCTCAAACTCACCT 607
|||
QY 710 GTGTAAAATTTAGCAATATATATAGACACAACCACTCCCAATGATCATCGTGGGTAAACAC 769
|||
Db 608 GTGTAAAATTTAGCAATATATATAGACACAACCACTCCCAATGATCATCGTGGGTAACTC 667
|||
QY 770 CTCCACACGAATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGTTACCTCAGCT 829
|||
Db 668 CTCCACACGAATAGTCTGCTACCTCAGGAATATTTTGTCTGTTGTTACCTCAGCT 727
|||
QY 830 ATCATTTGTTGAATGGCTCTTTCAGAACTATGCTCTCTCTCATTTCTAGTGCCCCCTA 889
|||
Db 728 ATCGTTGTTGAATGGCTC-TCAGAACTATGGCTCTCTCTCATTTCTTAGGGCCCCCTA 786
|||
QY 890 TGACCATCTACACTGAACAAGATT 913
|||
Db 787 TGACCATCTACACTGAACAAGATT 810
|||

RESULT 17
BX368078/c
LOCUS
DEFINITION
BX368078 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI051YM13 3-PRIME, mRNA sequence.
ACCESSION
BX368078
VERSION
BX368078.2 GI:46572908
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 846)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30447710.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
```

4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?S=CS0A012D05_U01092_1&c=4215.r.

FEATURES

Location/Qualifiers
1. .846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 47.2%; Score 698.6; DB 5; Length 846;
Best Local Similarity 92.8%; Pred. No. 2.5e-195;
Matches 786; Conservative 0; Mismatches 56; Indels 5; Gaps 5;

QY 467 AAACCTCCGTACCCATACCTCGCTGGTGAAGCCTATTTTAATACCAC-CCTCACTCGGCTC 525
DB 846 AAACCTCCGTACCCATACCTCGCTGGTGAAGCCTATTTTAATACCACCTCACTCGGCTC 787

QY 526 CATGAGGTCTACGCC-AAAACCTACTACTCTTGGATGTGCTCCCTCGACCTTCA- 583
DB 786 ATATAGGTCTCGGCCCAAAAACCTACTACTGTGAATATGCTCCCTCGAATTCAG 727

QY 584 GGCCATACATTTCAAT-CCCTGTTCTCGAACAATGGAACAATTTAGCAGCAGAAATAAAC 642
DB 726 GGCCATATGTTTCAATNCCCTGTACCTGANGCAATGNAACAACTTCAGCAGAAATAAAC 667

QY 643 ACCACTTCCGTTTGTAGTAGACCTCTTGTTCCTCAATCTGGAATTAACCCATACCTCAAC 702
DB 666 ACCACTTCCGTTTGTAGTAGACCTCTTGTTCCTCAATCTGGAATTAACCCATACCTCAAC 607

QY 703 CTACCTGTGTAAATTTAGCAATATATAGACACACAGCTCCCAATGCATCAGTGG 762
DB 606 CTCACCTGTGTAAATTTAGCAATATATAGACACACAGCTCCCAATGCATCAGTGG 547

QY 763 GTAACACTCCACAGCAATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGTACC 822
DB 546 GTAACCTCTCCACACAAATAGTCTGCTACCTCCCTCAGGAATATTTTGTCTGTGTACC 487

QY 823 TCAGCCTATCATTTGTTGAATGGCTTTTCAAGAACTTAATGTCTCTCTCATTTCTTAGTG 882
DB 486 TCAGCCTATCGTTGTTGAATGGCTTTTCAAGAACTTAATGTCTCTCTCATTTCTTAGTG 427

QY 883 CCCCCTATGACCATCTACCTGAACAAGATTTATACATCATGTCTGTACCTAAGCCCCAC 942
DB 426 CCCCCTATGACCATCTACCTGAACAAGATTTATACATCATGTCTGTACCTAAGCCCCAC 367

QY 943 AACAAAGAGTACCCATCTCTCTTTGTTATACAGACAGAGTGTAGGACAGACTAGGT 1002
DB 366 AACAAAGAGTACCCATCTCTCTTTGTTATAGGACAGAGTGTAGGACAGACTAGGT 307

QY 1003 ACTGGCAATGGCAGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAAATA 1062
DB 306 ACTGGCAATGGCAGTATCACAACTCTACTCAGTTCTTACTACAACTATCTCAAGAAATA 247

QY 1063 AATGGTGAATGGAAACAGGTACATGACTCCCTGTGTACCTTGGAAAGATCAACTTAATCC 1122
DB 246 AATGGGACATGGAAACGGGTCCGCACTCCCTGTGTACCTTGGAAAGATCAACTTAATCC 187

QY 1123 CTACGACGAGTGTCTTCAAAATCGAGAGCTTTAGACTGTCTAAACGCCCAAGAGGG 1182
DB 186 CTACGACGAGTGTCTTCAAAATCGAGAGCTTTAGACTGTCTTAAACGCCCTGAAAGAGGG 127

QY 1183 GGAACCTGTTTATTTTGGGAGAGAAACCGCTGTTTATTTATGTTAATCAATCAGAAATGTC 1242
DB 126 GGAACCTGTTTATTTTGGGAGAGAAATGCTGTTTATTTATGTTAATCAATCAGAAATGTC 67

QY 1243 ACTGAGAAGTTAAAGAAATTCGAGATCGAATACAAATCTAGACAGAGAGCTTCAAAAC 1302
DB 66 ACTGAGAAGTTAAAGAAATTCGAGATCGAATACAACTGAGCAGAGAGCTTCAAAAC 8

QY 1303 ACCGAAC 1309
DB 7 ATGGACC 1

RESULT 18

BX389657
LOCUS BX389657 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CS0DI051YM13 5-PRIME, mRNA sequence.
ACCESSION BX389657
VERSION BX389657.2 GI:46875146
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 995)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30462931.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?S=CS0BAF027ZB05_AF02531_2&c=4215.r

FEATURES

Location/Qualifiers
1. .995
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 47.0%; Score 696.2; DB 5; Length 995;
Best Local Similarity 86.8%; Pred. No. 1.4e-194;
Matches 826; Conservative 0; Mismatches 114; Indels 12; Gaps 5;

QY 266 GGAAATGATTAATCTCTAGTTCCTGGAGACTTGGAGCCTGTCTGTGGACTTACT 325
DB 1 GGAAATGATTAATCTCTAGTTCCTGGAGACTTGGAGT-CTGTCTGTGGACTTACT 59

QY 326 TCACCCATACAGTATCTCTGTATGGGGTGGAAATTCAGGTTCAGGCAGAGAGAAAAACAG 385
DB 60 TCACCCAAATCTGTATCTGTATGGGGTGGAGTTCAGATCAGCAGAGAAAAACATG 119

QY 386 TAAAGGAGCAATCTCCCACTGACCCGGGACATAGACCCCTAGCCCTTACAAAGAC 445
DB 120 TAAAGAGGTAATCTCCCACTCACCCTGGGTACATGGCACCTCTAGCCCTTACAAAGAC 179

QY 446 TAGTTCTCTCAAAACTACATGAAACCTTCGTACCCATCTCGCTCGTGGAGCTATTTA 505
DB 180 TAGATCTCTCAAAACTACATGAAACCTTCGTACCCATCTCGCTCGTGGAGCTATTTA 239


```

Db      831 AGCCATACTACATACACAAACCAATTCCTCATGCATCAAGGTGGTAACTCCTCCACAC-A 889
QY      781 ATAGTCTGCCTACCTCAGCAATATTTT 809
Db      890 ATAGTCTGCCTACCTCAGCAATATTTT 918

RESULT 20
LOCUS   AUI38405              723 bp  mRNA  linear  EST 02-AUG-2002
DEFINITION
AUI38405 PLACE1 Homo sapiens cDNA clone PLAC1008489 5', mRNA
sequence.
ACCESSION
AUI38405
VERSION
AUI38405.1 GI:10999926
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 723)
AUTHORS
Oka,T., Nishikawa,T., Suzuki,Y., Iehii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE
HRI human cDNA project
JOURNAL
Unpublished (2000)
COMMENT
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1..723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLAC1008489"
/tissue_type="placenta"
/clone_lib="PLAC1"
/notes="Vector: pME18SFL3"

ORIGIN
Query Match 43.7%; Score 647; DB 1; Length 723;
Best Local Similarity 94.6%; Pred. No. 4.9e-180;
Matches 679; Conservative 0; Mismatches 38; Indels 1; Gaps 1;

QY      565 TGCCTCCCCCTGCACCTTCAGCCCATACATTTCAATCCCTGTTCTGTGAACAATGGAACAAC 624
Db      3 TGCCTCCCCCTGAACTTCAGGCCATATGTTTCAATCCCTGTACCTGAACAATGGAACAAC 62
QY      625 TTACGACAGAAATAAACAACACTTCGGTTTGTAGGACCTCTGTGTTTCCAATCTGAA 684
Db      63 TTACGACAGAAATAAACAACACTTCGGTTTGTAGGACCTCTGTGTTTCCAATCTGAA 122
QY      685 ATACCCCATACCTCAAACTCACCTGTGTAAATTTAGCAATCTATAGACACAACACGAC 744
Db      123 ATACCCCATACCTCAAACTCACCTGTGTAAATTTAGCAATCTATAGACACAACACGAC 182
QY      745 TCCCAATGCATCAGGTGGGTAAACACTCCCAACAGAAATAGTCTGCCTACCCCTCAGGAATA 804
Db      183 TCCCAATGCATCAGGTGGGTAACTCTCCCAACAGAAATAGTCTGCCTACCCCTCAGGAATA 242
QY      805 TTTTGTGTGTGGTACCTCAGCCCTATCATTTGTTTGAATGGCTCTTTCAGAACTCTATGTGC 864
Db      243 TTTTGTGTGTGGTACCTCAGCCCTATCATTTGTTTGAATGGCTCTTTCAGAACTCTATGTGC 302
QY      865 TTCCTCTCAATCTTAGTGTCCCTCATGACCTACCTGACCTGAACAGATTTTATACAATCAT 924

```

```

Db      303 TTCCTCTCAATCTTAGTGCCCTATGACCATCTACACTGAACAAGATTTATACAGTTAT 362
QY      925 GTGCTACCTAAGCCCCCAACAAGAGTACCCATTTCTTCTTTTGTATCAGAGCAGGA 984
Db      363 GTCATATCTAAGCCCCGCAACAAGAGTACCCATTTCTTCTTTTGTATGAGAGCAGGA 422
QY      985 GTGCTAGGAGAGCTAGGTACTGGCATTTGGCAGTATCACAACCTCTACTCAGTTCTACTAC 1044
Db      423 GTGCTAGGTGCATAGGTACTGGCATTTGGCGGTATCACAACCTCTACTCAGTTCTACTAC 482
QY      1045 ARAACTATCTCAAGAAATAAATGTTGACATGGAGACAGTCACTGACTCCCTGGTCACTTG 1104
Db      483 AACTATCTCAAGAACTAATGGGGAACATGGAGACGGTGGCCGACTCCCTGGTCACTTG 542
QY      1105 CRAAGATCAAACTTAACCTCCCTAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTG 1164
Db      543 CRAAGATNAACTTAACCTCCCTAGCAGCAGTAGTCTTTCAAAATCGAAGAGCTTTAGACTG 602
QY      1165 CTAACGCCCAAAAGAGGGGGAACCTGTTATTTTAGGAGAGAACGCTGTTATTTATGTT 1224
Db      603 CTAACCGCTGAAAGANGGGGAACCTGTTAAATTTTAGGGGAAGAATGCTGTTATTANGTT 662
QY      1225 AATCAATCCAGAAATTCCTCACT-CAGAAAGTTAAAGAAATTCGAGATCGAATAACAATCT 1281
Db      663 AATCAATCCGGGATCGTCACTGGAGAAAGTTAAAGAAATTCGAGATCGAATAACAATCT 720

RESULT 21
LOCUS   AQ261133              631 bp  DNA  linear  GSS 24-OCT-1998
DEFINITION
CITBI-E1-2506C15.TF CITBI-E1 Homo sapiens genomic clone 2506C15,
genomic survey sequence.
ACCESSION
AQ261133
VERSION
AQ261133.1 GI:3787657
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 631)
AUTHORS
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL
Unpublished (1998)
COMMENT
Other_GSSs: CITBI-E1-2506C15.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
source
1..631
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2506C15"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBI-E1"
/notes="Vector: pBelosAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

ORIGIN
Query Match 42.5%; Score 629.4; DB 8; Length 631;

```


Best Local Similarity 99.8%; Pred. No. 7.8e-175;
Matches 630; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 378 AAACAAGTAAGGAGCAATCTCCAACTGACCCGGGACATACAGCCCTAGCCCTA 437
Db 1 AAACAAGTAAGGAGCAATCTCCAACTGACCCGGGACATACAGCCCTAGCCCTA 60

QY 438 CAAGGACTAGTCTCTCAAACTACATGAACCTCCGTATCCCACTACTCGCTGCTGAG 497
Db 61 CAAGGACTAGTCTCTCAAACTACATGAACCTCCGTATCCCACTACTCGCTGCTGAG 120

QY 498 CTAATTTATACACCTCCTCAGCTCCATGAGTCTCAGCCCAAAACCTTACTAATG 557
Db 121 CTAATTTATACACCTCCTCAGCTCCATGAGTCTCAGCCCAAAACCTTACTAATG 180

QY 558 TTGGATGTCCTCCCTGCCTCAGGACATATCTCAATCCCTGCTTGAACAATG 617
Db 181 TTGGATGTCCTCCCTGCCTCAGGACATATCTCAATCCCTGCTTGAACAATG 240

QY 618 GAACAACCTTCAGACAGAAATAAACACACTTCCGTTTTAGTAGACCTCTTGTTCCTCA 677
Db 241 GAACAACCTTCAGACAGAAATAAACACACTTCCGTTTTAGTAGACCTCTTGTTCCTCA 300

QY 678 TCTGGAATTAACCTTACCTCAACCTCCTGTTGTAAATTTAGCAATATATAGACAC 737
Db 301 TCTGGAATTAACCTTACCTCAACCTCCTGTTGTAAATTTAGCAATATATAGACAC 360

QY 738 AACAGCTCCCAATGATCAGTGGGTAAACCTCCACACAGAAATGCTGCTACCTC 797
Db 361 AACAGCTCCCAATGATCAGTGGGTAAACCTCCACACAGAAATGCTGCTACCTC 420

QY 798 AGGAATATTTTGTCTGTTGCTGCTCAGCTTACCTTATGTTGTTGATGCTTTCAGAATC 857
Db 421 AGGAATATTTTGTCTGTTGCTGCTCAGCTTACCTTATGTTGTTGATGCTTTCAGAATC 480

QY 858 TATGCTCTTCTCTCATTTCTAGTGCCTCTATGACCATCTACACTGACCAAGATTATA 917
Db 481 TATGCTCTTCTCTCATTTCTAGTGCCTCTATGACCATCTACACTGACCAAGATTATA 540

QY 918 CAATCATGTCGTAACCTAAGCCCAACAAAGAGTACCCATCTCTCTTTGTTATCAG 977
Db 541 CAATCATGTCGTAACCTAAGCCCAACAAAGAGTACCCATCTCTCTTTGTTATCAG 600

QY 978 AGCAGAGTCTAGGACAGACTAGGTAGTGGC 1008
Db 601 AGCAGAGTCTAGGACAGACTAGGTAGTGGC 631

RESULT 22
LOCUS AU138097 702 bp mRNA linear EST 02-AUG-2002
DEFINITION AU138097 PLACE1 Homo sapiens cDNA clone PLACE1007839 5', mRNA
sequence.
ACCESSION AU138097
VERSION AU138097.1 GI:10999618
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 702)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
Location/Qualifiers
1. 702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1007839"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Vector: pME18SFL3"

FEATURES
source

ORIGIN

Query Match 42.4%; Score 628.6; DB 1; Length 702;
Best Local Similarity 94.6%; Pred. No. 1.4e-174; Indels 1; Gaps 1;
Matches 660; Conservative 0; Mismatches 37;

QY 633 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATCTGGAATAAACCCA 692
Db 1 AGAAATAAACACCACTTCCGTTTTAGTAGGACCTCTTGTTCCTCAATCTGGAATAAACCCA 60

QY 693 TACCTCAAACTCCTACCTGTGTAAATTTAGCAATATATAGACACCAACCTCCCAATG 752
Db 61 TACCTCAAACTCCTACCTGTGTAAATTTAGCAATATATAGACACCAACCTCCCAATG 120

QY 753 CATCAGTGGGTAAACCTCCACAGCAATAGTCTGCTACCTCAGCAATATTTTGT 812
Db 121 CATCAGTGGGTAACTCTCTCCACACAAATAGTCTGCTACCTCAGCAATATTTTGT 180

QY 813 CTGTGGTACCTCAGCTATCATTTGTTGAATGGCTCTTCAGAACTCTATGCTCTCTC 872
Db 181 CTGTGGTACCTCAGCTATCATTTGTTGAATGGCTCTTCAGAACTCTATGCTCTCTC 240

QY 873 ATTCTTAGTGCCTCTATGACCATCTACCTGAAACAAGATTATACATCATGCTGCTACC 932
Db 241 ATTCTTAGTGCCTCTATGACCATCTACCTGAAACAATATTTATACATCATGCTGCTACC 300

QY 933 TAAGCCCCCAACAAAGAGTACCCATCTCTTCTTGGTTATCAGAGCAGGAGTCTAGG 992
Db 301 TAAGCCCCCAACAAAGAGTACCCATCTCTTCTTGGTTATCAGAGCAGGAGTCTAGG 360

QY 993 CAGACTAGTACTGGCATTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTATC 1052
Db 361 TGCAGTACTGGCATTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTATC 420

QY 1053 TCAAGAAATAAATGGTGACATGGAAACAGGTCACTGCTCCCTGGTCACTTGCAGATCA 1112
Db 421 TCAAGAAATAAATGGTGACATGGAAACAGGTCACTGCTCCCTGGTCACTTGCAGATCA 480

QY 1113 ACTTAACTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGTAAACCGC 1172
Db 481 ACTTAACTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTTTAGACTTGTAAACCGC 540

QY 1173 CAAAAGAGGGGAACCTGTTTATTTTAGGAGAGAAGACGCTGTTTATTTATTAATCAATC 1232
Db 541 TGAAGAGGGGAACCTGTTTATTTTAGGAGAGAAGATGCTGTTTATTTATTAATCAATC 600

QY 1233 CAGAATTGCTACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATAGTAG-AGCAGAGG 1291
Db 601 CGGAATCTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATACAAATAGTAG-AGCAGAGG 660

QY 1292 AGCTTCAAAACACCGAAGCTGGGCTCCCTCAGCCAA 1329
Db 661 AGCTTCAAAACACCGAAGCTGGGCTCCCTCAGCCAA 698

RESULT 23
LOCUS BX409328
DEFINITION BX409328 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE0131120 5-PRIME, mRNA sequence.

1104 GCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGAC 1161
 |||
 58 GCAAGATCAACTTAACCTCCCTAGCAGCAGTAGTCCTTCAAAATCGAAGAGCTTTAGAC 1
 |||

RESULT 26
 BX409304
 LOCUS
 DEFINITION
 5-PRIME, mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

186 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 792)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 15, 2003 this sequence version replaced gi:30766967.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS08AB0112H12_AE01057_1&c=4215.r

FEATURES
 source
 1..792
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS08012YJ24"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN
 Query Match 37.5%; Score 555.8; DB 5; Length 792;
 Best Local Similarity 88.0%; Pred. No. 5.7e-153;
 Matches 610; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
 29 TTACTGTTCTTACCCCTTTCCTCTCACTGCACCCCTCCATGCTGCTGACACCA 88
 1 TTACTGTTCTTACCCCTTTCCTCTCACTGCACCCCTCCATGCTGCTGACCA 60
 89 GTAGTCTCCCTTACCAAGAGTTTCTATGAAGAAGCGGCTTCTCTGGAATATTGATGCC 148
 61 GTAGTCTCCCTTACCAAGAGTTTCTATGGAGATGCAGCGTCCCGGAATATTGATGCC 120
 149 CATCATATAGAGTTTATCTAAGGAACTCCACCTTCACTGCCACCACTATGCC 208
 121 CATCGTATAGAGTCTTCTAAGGAACTCCACCTTCACTGCCACCACTATGCC 180
 209 GCACTCTATACTCTGCCACTCTTTCATGATGATGAATATCTATTATGGACAGGA 268
 181 GCAACTGCTATCACTCTGCCACTCTTTCATGATGATGAATATCTATTATGGACAGAA 240
 269 AAATGATTATCTCTAGTTGCTCTGGAGCACTTGGAGCACTGTCTGTGGACTTACTTCA 328
 241 AAATGATTATCTCTAGTTGCTCTGGAGCACTTGGAGCACTGTCTGTGGACTTACTTCA 300

329 CCCATACCAGTATGTCTGATGGGGTGGAAATTCAGGTCAGCAAGAGAAAAACAAGTAA 388
 |||
 301 CCCAACTCGTATGTCTGATGGGGTGGAGTTCAAGATCAGGCAAGAGAAAAACATGAA 360
 |||

389 AGGAAGCAATCTCCAACTGACCCCGGACATAGCACCCCTACCAAGGACTAG 448
 |||
 361 AAGAAGNAATNTNCCAACCTAACCCCGNTACATCGGCACTTNTAGCCCTACAAGGACTAG 420
 |||

449 TTCTCTCAAACTACATGAACCCCTCCGTAACCATCTCGCTGGTGGAGCTTATTTAATA 508
 |||
 421 ATCTCTCAAACTACATGAACCCCTCCGNAACCATCTCGCTGGNAAGTCTTATTAATA 480
 |||

509 CCACCTCTCACTCGGCTCCATGAGGTCTCAGCCCAAAACCTACTAACTGTGTGATGTC 568
 |||
 481 CCACCTCTCACTCGCTCCATGAGGTTCGTTCCAAACCTACTACTACTGCTGTATATGCC 540
 |||

569 TCCTCTGCACTTCAGGCACTATCAATCTCTGCTTCTGACCAATGCAACCACTTCA 628
 |||
 541 TCCTCTGAACTTCAGCCCATATGCTCAATTCCTGTACTGAACACTGCAACACTTCA 600
 |||

629 GCACAGAAATAACA-CCACTTCCGTTTGTAGTAGGACCTCTTGTTCCTCAATCTGGAATA 687
 |||
 601 GCACAGTAACAATAACCCCTTCTGTTTGTAGTAGGACCTCTTGTTCCTCAATCTGTAATA 660
 |||

688 ACCATACCTCAAACTCACTGCTGTGTAATAATT 720
 |||

661 TCTCTTCTCAAACTTACTTACTGCTGTGTAATAATT 693
 |||

RESULT 27
 BX430055/c
 LOCUS
 DEFINITION
 3-PRIME, mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

1..814
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS08013Y120"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS08AB0022G11_CS00184_1&c=4215.r

FEATURES
 source
 1..814
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS08013Y120"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN
 Query Match 37.5%; Score 555.8; DB 5; Length 792;
 Best Local Similarity 88.0%; Pred. No. 5.7e-153;
 Matches 610; Conservative 0; Mismatches 82; Indels 1; Gaps 1;
 29 TTACTGTTCTTACCCCTTTCCTCTCACTGCACCCCTCCATGCTGCTGACACCA 88
 1 TTACTGTTCTTACCCCTTTCCTCTCACTGCACCCCTCCATGCTGCTGACCA 60
 89 GTAGTCTCCCTTACCAAGAGTTTCTATGAAGAAGCGGCTTCTCTGGAATATTGATGCC 148
 61 GTAGTCTCCCTTACCAAGAGTTTCTATGGAGATGCAGCGTCCCGGAATATTGATGCC 120
 149 CATCATATAGAGTTTATCTAAGGAACTCCACCTTCACTGCCACCACTATGCC 208
 121 CATCGTATAGAGTCTTCTAAGGAACTCCACCTTCACTGCCACCACTATGCC 180
 209 GCACTCTATACTCTGCCACTCTTTCATGATGATGAATATCTATTATGGACAGGA 268
 181 GCAACTGCTATCACTCTGCCACTCTTTCATGATGATGAATATCTATTATGGACAGAA 240
 269 AAATGATTATCTCTAGTTGCTCTGGAGCACTTGGAGCACTGTCTGTGGACTTACTTCA 328
 241 AAATGATTATCTCTAGTTGCTCTGGAGCACTTGGAGCACTGTCTGTGGACTTACTTCA 300

ORIGIN		ORGANISM	Homo sapiens
Query Match	36.3%;	Score	537.6; DB 5; Length 814;
Best Local Similarity	85.3%;	Pred. No.	1.4e-147;
Matches	671; Conservative	0; Mismatches	105; Indels 11; Gaps 6;
QY	379	AAACAGTAGTAAGGAAGCAATCTCCCAACTGACCCGGGACATAGACACCCCTAGACCCCTAC	438
DB	814	AAACCTGGAACTAGAAATTTCTCAACCCAGCGGAGCAAGGCACCTATACCCCTTA	755
QY	439	AAAGACTAGTTCTTCAAAACTACATGAACCCCTCCGTAACCATCTAGCGCTT-GGTGAG	497
DB	754	AAAGGTATAAACTGCAGAAAAACACTAAAAACCCCGGGGCCAATATAGCTGGGTAAG	695
QY	498	CTATTATTATACACCCCTCACTGGCTCCATGAGGTCTCAGCGCCCAAAACCTACTAACTG	557
DB	694	CCAAITTTAAACACCCCTCA--TGGGGCCCATAGGGCTCGGGCCAAAACCCCTACTAACTG	637
QY	558	TTGGATGTGCTCCCTCCCTGCACCTT---CAGGGCCATACATTTCAATCCCTGTTCCCTGAA-C	613
DB	636	TTGGATATGCTCCCTCCCTGNACTTCAAGGCCAAATGTTTCAATCCCTGTACTGNAACC	577
QY	614	AATGGAACAACCTT---CAGCAGAGAAATAAACACCACTTCCGTTTATAGGACCTCTTG	670
DB	576	AATGGANCAACTTTCAGCAGAGAAAAATAAACACCACTTCCGTTTATAGGACCTCTTG	517
QY	671	TTTCCAACTCTGGAAT-ACCCTATACCTCAAACTCACCCTGCTGTGTAATAATTTAGCAATCT	729
DB	516	TTTCCAACTCTGGAATTAACCACTACCTCAAACTCACCCTGCTGTGTAATAATTTAGCAATCT	457
QY	730	ATAGACAACAACGAGCTCCCAATGCATCAGGTGGGTAAACACCTCCCAACAGAAATAGTCTGC	789
DB	456	ACATACACAACAACCACTCCCAATGCATCAGGTGGGTAACTCCTCCCAACAGAAATAGTCTGC	397
QY	790	CTACCTCAGGAATATTTTTGTGTGTGTACTCTAGCCTATCATTTTGAATGGCTCT	849
DB	396	CTACCTCAGGAATATTTTTGTGTGTGTACTCTAGCCTATCATTTTGAATGGCTCT	337
QY	850	TCAGAACTATGTGCTTCTCTCATTTCTAGTGCCCTCATGACCATCTACACTCAACAA	909
DB	336	TCAGAACTATGTGCTTCTCTCATTTCTAGTGCCCTCATGACCATCTACACTCAACAA	277
QY	910	GATTATACAATCATGTCTAGCTAAGCCCCCAACAAGAGAGTACCCATTTCTCTTTT	969
DB	276	GATTATACAATCATGTCTAGCTAAGCCCCCAACAAGAGAGTACCCATTTCTCTTTT	217
QY	970	GTTATCAGACGAGGAGTGTAGCAGACTAGTACTGGCAATGGCAGTATCAACCTCT	1029
DB	216	GTTATAGGAGCAGGAGTGTAGGTCACCTAGGTACTGGCAATGGCAGTATCAACCTCT	157
QY	1030	ACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGAC	1089
DB	156	ACTCAGTTCTACTACAAACTATCTCAAGAAATAAATGGTGACATGGAACAGGTCACTGAC	97
QY	1090	TCCCTGGTCACCTTGCAGAGTCAACTTAATCCCTAGCAGCAGTGTAGTCTTCAAAATCGA	1149
DB	96	TCCCTGGTCACCTTGCAGAGTCAACTTAATCCCTAGCAGCAGTGTAGTCTTCAAAATCGA	37
QY	1150	AGAGCTT 1156	
DB	36	AGAGCTT 30	
RESULT 28			
LOCUS	BE019603		
DEFINITION	ba84f03.y1 NIH MGC_21 Homo sapiens cDNA clone IMAGE:2907101 5', similar to TR:O95244 O95244 ENVELOPE PROTEIN ; contains Alu repetitive element.; mRNA sequence.		
ACCESSION	BE019603		
VERSION	BE019603.1		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
ORGANISM			

```

QY 1183 GGAACCTGTTTATTTTAGGAGAAACGCTGTTATTTAATCAATCCAGAAATGTC 1242
Db 481 GGAACCTGTTTATTTTAGGAGAAACGCTGTTATTTAATCAATCCAGAAATGTC 540

QY 1243 ACTGAGAAATGTTAAAGAAATTCGAGATCGAATACAAATGAGAGCAG 1288
Db 541 ACTGAGAAATGTTAAAGAAATTCGAGATCGAATACAAATGAGAGCAG 586

RESULT 29
BG572445 790 bp mRNA linear EST 10-APR-2001
LOCUS 602593450F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4720846 5',
DEFINITION mRNA sequence.
ACCESSION BG572445
VERSION BG572445.1 GI:13580098
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 790)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: L1CM1576 row: c column: 23
High quality sequence stop: 624.
FEATURES
source Location/Qualifiers
1..790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4720846"
/lab_host="DH10B (Ti phage-resistant)"
/clone_lib="NIH_MGC_79"
/notes="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site 1: SfII (ggccgctcgcc); Site 2: SfiI
(ggccattatggc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGGCGGCGGGCGGCAGATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 36.1%; Score 535.2; DB 4; Length 790;
Best Local Similarity 90.5%; Pred. No. 7.3e-147;
Matches 650; Conservative 0; Mismatches 58; Indels 10; Gaps 7;

QY 742 AGCTCCCAATGATCAGTGGTGAACACCTCCACAGCAATAGTCTGCCTACCTCAGGA 801
Db 1 AACTCCCAATGATCAGTGGTGAACCTCCACAGCAATAGTCTGCCTACCTCAGGA 60

QY 802 ATATTTTGTGCTGTGTTACCTAGCTATCATTTGTTTGAATGCTCTTCAGATCTATG 861
Db 61 ATATTTTGTGCTGTGTTACCTAGCTATCATTTGTTTGAATGCTCTTCAGATCTATG 120

QY 862 TGCTTCCTCTCATTTCTTAGTGCCTCCCTACCATCTACACTGAACAGATTTTACAAT 921
Db 121 TGCTTCCTCTCATTTCTTAGTGCCTCCCTACCATCTACACTGAACAGATTTTACAAGT 180

```

```

QY 922 CATGTGCTAGTAAAGCCCAACAAAGAGTAGTACCATTCTTCCTTTGTTTATCAGAGCA 981
Db 181 TATGTATCATATCTAAGCCCGCAACAAAGAGTAGTACCATTCTTCCTTTGTTTATCAGAGCA 240

QY 982 GGAAGTCTAGGAGCAGACTAGGTACTGCGATTGGCAGTATCACAACCTCTACTCAGTCTTAC 1041
Db 241 GGAAGTCTAGGAGCAGACTAGGTACTGCGATTGGCAGTATCACAACCTCTACTCAGTCTTAC 300

QY 1042 TACAAACTATCTCAAGAAATAAATGTTGACATGGAACAGGTCACTGACTCCCTGGTCAACC 1101
Db 301 TACAAACTATCTCAAGAAATAAATGTTGACATGGAACAGGTCACTGACTCCCTGGTCAACC 360

QY 1102 TTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCA-AAATCGAAGAGCTTTAGA 1160
Db 361 TTGCAAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCA-AAATCGAAGAGCTTTAGA 420

QY 1161 CTTGCTTAACCGCCAAAGAGGGGGAACCTGTTTATTTTAGGAGAAAGAACGCTGTTTATTA 1220
Db 421 CTTGCTTAACCGCTGAAGAGGGGGAACCTGTTTATTTTAGGAGAAAGAACGCTGTTTATTA 480

QY 1221 TGTTAATCAATCCAGAAATTTGTCACTGAGAAAGTTA-AAGAAATTCGAGATCG--AAATACA 1277
Db 481 TGTTAATCAATCCGGAATCGTCACTGAGAAAGTTA-AAGAAATTCGAGATCGCAATACAC 540

QY 1278 ATGTAGAGCAGAGGAGCTTTCAAAACACCGAAGCTGGGGCCCTCTCAGCCAATGGATGCC 1337
Db 541 ACGTAGAGCAGAGGAGCTTTCAAAACACCTGCACCTCGGGG--TCCTCAGCAATGGATGCC 598

QY 1338 CTGGGTTCTCCCTTCTTAGGACCTCTAGC-AGCTCTAATATTTGTTACTCTCTTTGGAC 1396
Db 599 --TGGATTCTCCCTTCTTAGGACCTCTAGCAAGCTATATCTATTGCTACTCTCTC-TTGGAG 655

QY 1397 CCTGTATCTTTAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAATGGAAGCTGTAAGCT 1454
Db 656 CCGTATCTTTAACTCTCTTGTCACTGAGTCTCTTCCAGAAATCGAAGGTGTAAGCT 713

RESULT 30
BX450495 760 bp mRNA linear EST 05-MAY-2004
LOCUS BX450495 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX450495
VERSION BX450495.2 GI:47038997
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 760)
JOURNAL Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31018916.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0AU006ZG02_U0495_1kc=4215.r.
FEATURES
source Location/Qualifiers
1..760
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013Y120"
/tissue_type="PLACENTA"

```

/clone lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN	Query Match	32.3%;	Score 478.6;	DB 5;	Length 760;
	Best Local Similarity	83.3%;	Pred. No. 4.5e-130;		
	Matches 604;	Conservative 0;	Mismatches 114;	Indels 7;	Gaps 5;
Qy	682	GAATAAACCCATACCTCAAACTCCTACCTGTGTGAAAATTTAGCAATCT-ATAGACACAAC	740		
Db	727	GCACACCAAGCCCTCCCTCGGGGGGAAATCTCGTATCTCCCTATATCAAC	668		
Qy	741	CAGCTCCCAATGCATCAG---GTGGGTAAACACCTCCCAACAGATGTCTGCCTACCTCC	797		
Db	667	CGACCACCCATGCATCAGGTGAGTGGTCCCTCTCTCAAAATAGTTTGCTATACCTCC	608		
Qy	798	AGGAATATTTTGTCTGTGTGCTACCTCAGCCTATCATTTGTTGAATGCTCTTCAGATC	857		
Db	607	AGAAATAGTGGTGTGTGAGGCTCTCTCAGCCAAACCGTAGTATGAAGGGCTCTTCAGAAC	548		
Qy	858	TATGTCTTCCCTCTCATTTAGTGCCTCCCTATGACCATCTACACTGAAC-AAAGATTAT	916		
Db	547	AAATGTGC-TCCCTGGCAATCTAGAGCCCTATGACCATCGACATGAATTTAG	489		
Qy	917	ACAATCATGTCTGACTAAGCCCCCAACAAAAGAGTAGCCCAATCTTCTTTTGTATCA	976		
Db	488	CCCATATGTCTATCTAAGCCCGACCAACAAAAGAGTCCCATCTCTCTTTTGTATAG	429		
Qy	977	GAGCAGGAGTGTAGGACACTAGTACTGCGATTTGGCAGTATCAAA-CCTCTACTCAG	1035		
Db	428	AAGCAGGAGTGTAGGACACTAGGAACCTGGCAATTTGCGGTATCAAAAGCTCTACTCAG	369		
Qy	1036	TTTCTACTCAAACTACTCAAGAAATAAATGGTGACATGGAACAGGTCACTGACTCCCTG	1095		
Db	368	TTTCTACTCGGACTATCTCAAGACTAATATGGGACATGGAACGGGTCCCGACTCCCTG	309		
Qy	1096	GTCACTTCCAGAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGCT	1155		
Db	308	GTCACTTCCAGAGATCAACTTAACCTCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGCT	249		
Qy	1156	TTAGACTTCTACCCGCAAGAGGGGAACTGTTATTTTATTTAGGAGAAAGCGCTGT	1215		
Db	248	TTAGACTTCTACCCGCAAGAGGGGGAAGCTGTTATTTTATTTAGGAGAAAGTGTGT	189		
Qy	1216	TATTATGTTAAATCAATCCAGAAATTTGCTACTGAGAAAGTTAAAGAAATTCGAGATCGAATA	1275		
Db	188	TACTATGTTAAATCAATCCGGAATCTGCTACTGAGAAAGTTAAAGAAATTCGAGATCGAATA	129		
Qy	1276	CAATGTAGACGAGAGAGCTTCAAAACACCGAACGCTGGGGCTCTCTAGCAATGGATG	1335		
Db	128	CAACGTAGACGAGAGAGCTTCAAAACACCTGGACCTGGGGCTCTCTAGCAATGGATG	69		
Qy	1336	CCCTGGGTTCTCCCTTCTTAGACCTCTAGCAGCTCTAATATTTGTTACTCTCTTGA	1395		
Db	68	CCCTGGGTTCTCCCTTCTTAGACCTCTAGCAGCTCTAATATTTGTTACTCTCTTGA	9		
Qy	1396	CCCTG 1400			
Db	8	CCCTG 4			

RESULT 31
CR735592
LOCUS
DEFINITION
CR735592 NIH MGC 21 Homo sapiens cDNA clone IMAGE3840572 5', mRNA sequence.
ACCESSION
CR735592
VERSION
CR735592.1 GI:51584854
KEYWORDS
EST.

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 554); Ebert,L., Heil,O., Hennig,S., Korn,B., Neubert,P., Partsch,E., Peters,M., Radelof,U. and Schneider,D.
TITLE	I.M.A.G.E. cDNA Clone Collection
JOURNAL	Unpublished (2004)
COMMENT	Contact: Inge Axlart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Email: www.rzpd.de RZPD; IMAGE3840572.1 RZPDLIB; I.M.A.G.E. cDNA Clone Collection; Contact: Inge Axlart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 100 Fax: +49 30 32639 111 www.rzpd.de This clone is available royalty-free from RZPD; contact RZPD (Clone@rzpd.de) for further information. Seq primer: M13r, Primer sequence: TTTCCACAGGAACAGCTAIGAC. Location/Qualifiers 1..554 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE3840572" /tissue_type="choriocarcinoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 21" /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
FEATURES	source
ORIGIN	Query Match 32.1%; Score 475.4; DB 7; Length 554; Best Local Similarity 95.0%; Pred. No. 3.6e-129; Matches 491; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy	702 CCTCACCTGTGTAATAATTTAGCAATATATATAGACACACACCACTCCCAATGCATCAGGTG 761
Db	38 CCTCACCTGTGTAATAATTTAGCAATATATATAGACACACCACTCCCAATGCATCAGGTG 97
Qy	762 GGTAAACACCTCCACAGATAGTCTGCTACCCCTCAGGAATATTTTTTGTCTGTGGTAC 821
Db	98 GGTAACTCTCTCCACACAAATAGTCTGCTTACCCTCAGGAATATTTTTTGTCTGTGGTAC 157
Qy	822 CTCAGCCTTATCATTTGTAATGGCTCTTCAGAAATCTATGTGCTTCTCTCTCATTTCTAGT 881
Db	158 CTCAGCCTTATCGTTGTTGTAATGGCTCTTCAGAAATCTATGTGCTTCTCTCTCATTTCTAGT 217
Qy	882 GCCCCTATGACCATCTACACTGAACAGATTTATACAAATCATGTGCTGATCAAGCCCCA 941
Db	218 GCCCCTATGACCATCTACACTGAACAGATTTATACAAATCATGTGATCATATCTAAGCCCCG 277
Qy	942 CAACAAAGAGTACCCATCTCTCTTTGTTTATCAGACGAGAGTGTAGGCGACTAGG 1001
Db	278 CAACAAAGAGTACCCATCTCTCTTTGTTTATAGGACGAGAGTGTAGGCGACTAGG 337
Qy	1002 TACTGGCATTGGCAGTATCAACACTCTACCTCAGTTCTTACTACAAATCTCTCAAGAAAT 1061
Db	338 TACTGGCATTGGCGGTATCAACACTCTACTCAGTTCTTACTACAAATCTCTCAAGAACT 397
Qy	1062 AAATGGTGACATGGAACAGGTCACTGACTCCCTGGTCACTCTTGGAAATCAACTTAATC 1121


```

Db      398 AAATGGGACATGGAACGGTCCGACCTCCCTGGTCACTTGCAGATCAACTTAACTC 457
Qy      1122 CCTAGCAGAGTAGTCTCTTCAAAATCGAAGAGTTTGTAGACTTCCTAACGCCCAAGAGG 1181
Db      458 CCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGTTTGTAGACTTCCTAACGCCCAAGAGG 517
Qy      1182 GGGAACTGTTTATTTTATTTAGGAGGAAGAGCTGTTAT 1218
Db      518 GGGAACTGTTTATTTTATTTAGGGAAGAAGTCTGTTAT 554

RESULT 32
BX365066/c      1071 bp      mRNA      linear      EST 08-APR-2004
LOCUS      BX365066 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      Clone CS0D1046YA18 3-PRIME, mRNA sequence.
ACCESSION      BX365066
VERSION      BX365066.2 GI:46304105
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 1071)
JOURNAL      Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
COMMENT      Full-length cDNA libraries and normalization
On May 5, 2003 this sequence version replaced gi:30374869.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSIA1012ZC10NP1&c=4215.r.
FEATURES
source      1..1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized."
ORIGIN
Query Match      30.0%; Score 444; DB 5; Length 1071;
Best Local Similarity      89.5%; Pred. No. 9.2e-120;
Matches 485; Conservative      3; Mismatches 53; Indels 1; Gaps 1;

Qy      940 CACAACAAAAGAGTACCATTCTTCTTTTGTATCATGAGCAGGAGTGCTAGGCAGACTA 999
Db      1026 CCCSCACAAAAGAGWCCATCTCTCTTCTTTATAGGAGCAGAGTGCTA-GTGCACTA 968
Qy      1000 GGTACTGGCATTTGGCAGTATCACAACTTACTCAGTTCTTACTACAACATCTCTCAGAA 1059
Db      967 GGGACTGGCATTTGGCGGTATACAACTCTTACTCAGTTCTTACTACAACATCTCTCAGAA 908
Qy      1060 ATAAATGGTGACATGGACAGGTGACCTGACTCCCTGTGCTCCTTGCAGATCACTTAAAC 1119
Db      907 CTAATGGGGACATGGAAACGGGTGCGGACCTCCCTGGTCACTTGCAGATCACTTAAAC 848
Qy      1120 TCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTTGCTTAACGCCCAAGA 1179
Db      847 TCCCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTTGCTTAACGCCGTGAAGA 788

```

```

Qy      1180 GGGGAACCTGTTTATTTTATTTAGGAGAGAACCTGTTTATTTATGTTAATCAATCCGAAT 1239
Db      787 GGGGAACCTGTTTATTTTATTTAGGGAAGAAATGCTGTTTATTTATGTTAATCAATCCGAATC 728
Qy      1240 GTCACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACATATGATGAGCAGAGGAGCTTCAA 1299
Db      727 GTCACTGAGAAAGTTTAAAGAAATTCGAGATCGAATACATATGATGAGCAGAGGAGCTTCCA 668
Qy      1300 AACACCGAACGCTGGGGCTCTCTCAGCCAAATGATGCCCTGGTGTCTCCCTCTTTTAGGA 1359
Db      667 AACACTGGACCTGGGGCTCTCTCAGCCAAATGATGCCCTGGTGTCTCCCTCTTTTAGGA 608
Qy      1360 CCTCTAGCAGCTCTAATATTTGTTACTCTCTTTTGGACCTGATCTTTAACTCTCTTGT 1419
Db      607 CCTCTAGCAGCTATAATATTTGTTACTCTCTTTTGGACCTGATCTTTAACTCTCTTGT 548
Qy      1420 AAGTTTGTCTCTCCAGATTCGAAGCTGTAAGCTTACAGATGCTTTACAAATGGAACCC 1479
Db      547 AACTTTGTCTCTTCCAGATTCGAAGCTGTAAGCTTACAAATGGAACCCAGATGCGAGTCC 488
Qy      1480 CA 1481
Db      487 AA 486

RESULT 33
BX368759
LOCUS      BX368759 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      clone CS0D1022YJ18 5-PRIME, mRNA sequence.
ACCESSION      BX368759
VERSION      BX368759.2 GI:46833898
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL      Full-length cDNA libraries and normalization
COMMENT      On May 8, 2003 this sequence version replaced gi:30459619.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAB021ZG07_AE01976_1&c=4215.r
FEATURES
source      1..944
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="CS0D1022YJ18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized."
ORIGIN
Query Match      28.5%; Score 421.6; DB 5; Length 944;
Best Local Similarity      81.8%; Pred. No. 3.9e-113;
Matches 535; Conservative      0; Mismatches 114; Indels 5; Gaps 4;

```



```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 998)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30337641.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODI0151AG07NP1&c=4215.r.
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI051M13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match 27.8%; Score 412.2; DB 5; Length 998;
Best Local Similarity 90.5%; Pred. No. 2.4e-110;
Matches 457; Conservative 4; Mismatches 42; Indels 2; Gaps 2;

QY 979 GCAGGAGTGTAGGCAGACTAGGTACTGGCAATGGCAGTATCAACC-TCTACTCAGTT 1037
Db 998 GCAGGAGTGTAGGTCAGTCTAGGACTGKCAATGAGTATCAACCTTCWCCTCAGTT 939
QY 1038 CTACTACAACTATCTCAAGAAAT-AAATGTGACATGGAACAGGTCACTGCTCCCTGG 1096
Db 938 CTACTACAACTATCTCAAGAACTTAAATGGGACATGGAACGGGTGCGGACTCCCTGG 879
QY 1097 TCACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTT 1156
Db 878 TCACCTTGCAGATCAACTTAACTCCCTAGCAGCAGTAGTCTTCAAAATCGAAGAGCTT 819
QY 1157 TAGACTTGTCTAACCGCCAAAGAGGGGAACTGTTTATTTTATAGGAGAAAGCGTGT 1216
Db 818 TAGACTTGTCTAACCGCTGAAAGAGGGGAACTGTTTATTTTATAGGAGAAAGCGTGT 759
QY 1217 ATTAGTTAATCAATCCAGAAATTTCTAGGAAAGTTAAAGAAATTCGAGATCAATAC 1276
Db 758 ATTAGTTAATCAATCCGGAATCTCACTGAGAAAGTTAAAGAAATTCGAGATCAATAC 699
QY 1277 AATGTAGCAGAGAGGCTTCAAAACACCGACCTGGGCTCTCAGCCAAATGGATGC 1336
Db 698 AATGTAGCAGAGAGGCTTCAAAACACCGACCTGGGCTCTCAGCCAAATGGATGC 639
QY 1337 CCTGGGTTCTCCCTTCTTAGGACTCTAGCAGCTCTAATATTTGTTACTCTCTTTGGAC 1396
Db 638 CCTGGATCTCCCTTCTTAGGACTCTAGCAGCTATAATATTTGTTACTCTCTTTGGAC 579
QY 1397 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCAGAAATTCGAGCTGTAAGCTAC 1456
Db 578 CCTGTATCTTTAACTCTCTTTAGTTTGTCTCTTCAGAAATTCGAGCTGTAAGCTAC 519
QY 1457 AGATGGTCTTACAATGAACCCCA 1481

Db 518 AAATGGAGCCCAAGATGCAGTCCAA 494

RESULT 36
BX378303/c
LOCUS
DEFINITION
BX378303 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI018YH07 3-PRIME, mRNA sequence.
ACCESSION
BX378303
VERSION
BX378303.2 GI:46557492
KEYWORDS
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1058)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30439129.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODI018CD04NP1&c=4215.r.
Location/Qualifiers
1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI018YH07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match 27.8%; Score 412.2; DB 5; Length 1058;
Best Local Similarity 87.9%; Pred. No. 2.5e-110;
Matches 457; Conservative 11; Mismatches 49; Indels 3; Gaps 2;

QY 962 TTCCTTTTGTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGGCATTGGCAGTATCA 1021
Db 1007 TTCCTTTTGTATCAGAGCAGGAGTCTAGGCAGACTAGGTACTGGCATTGGC-GTATCA 951
QY 1022 CAACCTCTACTCAGTCTTACTACAACTATCTCAAGAAATTAATGTTGATGAAACAGG 1081
Db 950 CAACCTCTACTCAGTCTTACTACAACTATCTCAAGAACTTAATGTTGATGAAACAGG 891
QY 1082 TCACCTGCTCCCTGGTCACTTGCAGATCACTTAATCTCCTAGCAGCAGTACTCCTTC 1141
Db 890 TCACCTGCTCCCTGGTCACTTGCAGATCACTTAATCTCCTAGCAGCAGTACTCCTTC 831
QY 1142 AAATCGAAGCTTTTGTAGCTTGTCTAACCGCCAAAGAGGGGAACTGTTTATTTTAG 1201
Db 830 AAATCGAAGCTTTTGTAGCTTGTCTAACCGCTGAAAGAGGKXKAWCYTTTTTTTAG 771
QY 1202 GAGAAGAACCTGTTTATTTATTTAATCAATCCAGAAATTTGCTACTGAGAAAGTTAAAGAA 1261
Db 770 GGAAGAAATCTGTTACTATGTTAATCAATCCKAWTCKTCACTGAGAAAGTTAAAGAA 711
QY 1262 TTCAGATCGAATACAAATGTAGAGCAGAGCTTCAAAACACCGAAGCTGGGGCTCC 1321

```

Db 710 TTGAGATCGAATACAACTAGACAGAGAGCTTCGAAACACTGGACCCCTGGGCTCC 651
 QY 1322 TAGCCAAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTCTAATATTGT 1381
 Db 650 TCAGCCAAATGGATGCCCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATTGC 591
 QY 1382 TACTCTCTTTGGACCTGTATCTTTAACTCTCTTTAACTTTGTTAACTTTGTTCTTCAGAAATG 1441
 Db 590 TACTCTCTTTGGACCTGTATCTTTAACTCTCTTTAACTTTGTTAACTTTGTTCTTCAGAAATG 531
 QY 1442 AAGCTGTAAAGCTACAGATGGTCTTTACAAATGGAACCCCA 1481
 Db 530 AAGCTGTAAAGCTACAAATGGAGCCCAAGATGCAGTCCAA 491

RESULT 37
 BX429316
 LOCUS
 DEFINITION BX429316 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120
 5-PRIME mRNA sequence.
 ACCESSION BX429316
 VERSION BX429316.2 GI:47035285
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 905)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 15, 2003 this sequence version replaced gi:30786832.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0AT011ZF02_T0974_1&c=4215.r.

FEATURES
 source
 1. .905
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE013Y120"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 Note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN
 Query Match 27.8%; Score 411.4; DB 5; Length 905;
 Best Local Similarity 82.5%; Pred. No. 4.1e-110;
 Matches 572; Conservative 0; Mismatches 101; Indels 20; Gaps 8;

QY 1 ATGGCCCTCCCTATCATATCTTTCTTTACTGTCTCTTACCCCTTCCTCTCACT 60
 Db 196 ATGGCCCTCCCTATCATATCTTTCTTTACTGTCTTTTACCCCTTTTCACTCTCACT 255
 QY 61 GCACCCCTCCCTATCATATCTTTCTTTACTGTCTCTTACCAAGAGTTCTTATGAGA 120
 Db 256 GCACCCCTCCCTATCATATCTTTCTTTACTGTCTTTTACCCCTTTTCACTCTCACT 315
 QY 121 ACGCGCTTCCTGGAAATATTGATGCCCATATAGGAGTTTATCTAAGGAAACTCC 180

Db * 316 ATGAGCGTCCGGAAATATTGATGCCCATCTGATAGGAGTCTTTTAAAGGGAACCCCC 375
 QY 181 ACCTTCACTGCCACACCCATATGCCCCGCAACTGCTATACTCTGCCACTCTTTTGCAATG 240
 Db 376 ACCTTCACTGCCACACCCCATGCCCCGCAACTGCTATACTCTGCCACTCTTTTGCAATG 435
 QY 241 CATGCAAACTACTCATTATTGGACAGGGAATGATTAAATCCTAGTTGCTCTGGAGACTT 300
 Db 436 CATGCAAACTACTCATTATTGGACAGGGAATGATTAAATCCTAGTTGCTCTGGAGACTT 495
 QY 301 GGAGCACTGTCTGTGGACTTACTTACCATACCAGTATGTCTGTATGGGGTGGAAAT 360
 Db 496 GGAGTCACTGTCTGTGGACTTACTTACCATACCAGTATGTCTGTATGGGGTGGAGTT 555
 QY 361 CAAGTCAAGCAAGAGAAACAGTAAGGAAGCAATCTCCCACTGACCCGGGAGCAT 420
 Db 556 CAAGTCAAGCAAGAGAAACAGTAAGGAAGCAATCTCCCACTGACCCGGGAGCAT 615
 QY 421 AGCACCCTAGCCCCCTACAAAGGACTAGTTTC-TCTCAAACTACATG--AAACCTCCGT 477
 Db 616 GGCACCTCTAGCCCCCTACAAAGGACTAGAACTTCTCAAACTACATGGAACCCCTCCGT 675
 QY 478 ACCC--ATACTCGCTGTGTAGCCCTATTATACACCCCTCACTCGGT-----CCA 527
 Db 676 ACCCATTTACTCAGCTTGGTAAAGCCTATTTTAATACCCACCCCTGACTGGGGCTTCCAT 735
 QY 528 TGAGGTCTCAGCCCAAAACCCCT--ACTAACTGTTGG---ATGTGCTCCCCCTGCACCTTC 582
 Db 736 GAAGGTCTAGGACCCCAAGCCCTTAACTAACTGTGTGGATATGCCCTTCCCCTGAGCTTTA 795
 QY 583 AGGCATACATTTCAATCCCTGTTCTTGAACAATGGAACAAC-TTCAGCAGCAAAATAAA 641
 Db 796 AGCCCTATGTTGCAATGCCGTTCCTATACAATGAAGCAACTTTTATCCCAAGTAATAA 855
 QY 642 CACAC-TTCGCTTTTAGTAGGACCTCTTGT 673
 Db 856 CCCCACTTTCGCTTTTAGTAGGACCTTGTGT 888

RESULT 38
 BX388917/c
 LOCUS
 DEFINITION BX388917 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1022YJ18 3-PRIME, mRNA sequence.
 ACCESSION BX388917
 VERSION BX388917.2 GI:46845081
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 564)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 8, 2003 this sequence version replaced gi:30461045.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0BAB005ZC06_CS00428_1&c=4215.r

FEATURES
 source
 1. .564
 /organism="Homo sapiens"
 /mol_type="mRNA"

/db xref="taxon:9606"
/clone="CS0DI022YJ18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 27.4%; Score 406.4; DB 5; Length 564;
Best Local Similarity 84.4%; Pred. No. 1.1e-108;
Matches 467; Conservative 0; Mismatches 84; Indels 2; Gaps 1;
QY 813 CTGTGTACCTGACCTATCATTTGTTGATGCTCTTCAGATCTATGTGCTTCTCTC 872
DB 564 CTGGGAACCAACCAACAGCATGTTAAGAGACAGTAAAGAAACGAGAGCTACACAG 505
QY 873 ATTCTTAGTCCGCCCTAT--GACCATCTACACTGAACAGATTATACATCATGTCGTA 930
DB 504 ATACTAAGAGAGCTATGAGACAAACGCTGGAGCAAGATTATCATGTTATGTCATA 445
QY 931 CCTAAGCCCCACAAACAGAGTACCCATTCTCTTTCTTTCTATTCAGAGCAGAGTGCTA 990
DB 444 TTTAAGCCCCGCATCAAAAGAGTACCCATTCTCTTTCTTTCTATAGGAGCAGAGTGCTA 385
QY 991 GGCAGACTAGGTACTGGCATTGGCAGTATCACAACTCTACTCAGTTCTACTACAACTA 1050
DB 384 GGTGCATAGGTACTGGCATTGGCGGTATCACAACTCTACTCAGTTCTACTACAACTA 325
QY 1051 TCTCAAGAAATAATGTGACATGGAACAGTCACTGACCTCCCTGGTCACTTGGCAAGAT 1110
DB 324 TCTCAAGAAATAATGGGACATGGAACGGTGGCCGACCTCCCTGGTCACTTGGCAAGAT 265
QY 1111 CAACCTTAACCTCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTAAAC 1170
DB 264 CAACCTTAACCTCTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTAAAC 205
QY 1171 GCCAAAGAGGGGAACCTGTTATTTTAGGAGAGAAACGCTGTTATTTATGTTAAATCAA 1230
DB 204 GCTGAAGAGGGGAACCTGTTATTTTAGGAGAGAAACGCTGTTATTTATGTTAAATCAA 145
QY 1231 TCCAGATTGTCTAGTGAAGAGTTAAAGAAATTCGAGATCGAATACAAATGTAGAGCAGAG 1290
DB 144 TCCGGAATGTCTAGTGAAGAGTTAAAGAAATTCGAGATCGAATACAACTAGAGCAGAG 85
QY 1291 GAGCTTCAAAACACCGAACGCTGGGCGCTCTCTAGCCAAATGGATGCCCTGGGTTCTCCCC 1350
DB 84 GAGCTTCAAAACACCTGGACCTGGGCGCTCTCTAGCCAAATGGATGCCCTGGATTGNCNCCC 25
QY 1351 TTCTTAGAGACCTC 1363
DB 24 CTNCTAGAGACCCC 12

RESULT 39
BX439636/c
LOCUS
DEFINITION BX439636 Homo sapiens PLACENTA COT 25-NORMALIZED EST 04-MAY-2004
3-PRIME, mRNA sequence.
ACCESSION BX439636
VERSION BX439636.2 GI:47000005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1019)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30771765.
Contact: Genoscope

Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.1
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DE012DE12NP1&c=4215.1.

FEATURES

source
1..1019
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 27.1%; Score 402; DB 5; Length 1019;
Best Local Similarity 90.0%; Pred. No. 2.6e-107;
Matches 434; Conservative 6; Mismatches 41; Indels 1; Gaps 1;
QY 1001 GTACTGGCATTGCGAGTATCAACCTCTACTCAGTTCTACTCAAACTATCTCAAGAAA 1060
DB 975 GGWAYKGCATTGGCGTWTMTTAAACCACTCTACTCAGTTCTACTCAAACTATCTCAAGAAC 916
QY 1061 TAAATGGTGCATGAGCAAGTCACTGACCTCCCTGGTCACTT-GCAAGATCAACTTAAC 1119
DB 915 TAAATGGGACATGGAACCGGTCGCGACTCCCTGGWCACCTTGGCAAGATCAACTTAAC 856
QY 1120 TCCTTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAAACGCAAAAGA 1179
DB 855 TCCTTAGCAGCAGTAGTCTCTTCAAAATCGAAGAGCTTTAGACTTGTCTAAACGCAAAAGA 796
QY 1180 GGGGGAACCTGTTATTTTAGGAGAAACCGTCTTATTTATGTTATCAATCCAGATT 1239
DB 795 GGGGGAACCTGTTATTTTAGGAGAAATGCTGTTACTATGTTTAACTAATCCGAATC 736
QY 1240 GTCACTGAGAAAGCTTAAAGAAATTCGAGATCGAATCAATGTAGAGCAGAGGAGCTTCAA 1299
DB 735 GTCACTGAGAAAGCTTAAAGAAATTCGAGATCGAATCAATGTAGAGCAGAGGAGCTTCAA 676
QY 1300 AACACCGAACGCTGGGCGCTCTCTAGCCAAATGGATGCCCTGGGTTCTCCCTTCTTAGGA 1359
DB 675 AACCTGGACCCCTGGGCGCTCTCTAGCCAAATGGATGCCCTGGATTTCTCCCTTCTTAGGA 616
QY 1360 CCTCTAGCAGCTCTAATATTGTTACTCTCTTTGGACCCCTGATCTTTTAACTCTCTGTT 1419
DB 615 CCTCTAGCAGCTCTAATATTGTTACTCTCTTTGGACCCCTGATCTTTTAACTCTCTGTT 556
QY 1420 AAGTTTGTCTCTCCAGAAATTAAGCTTAAGCTACAGATGGTCTTACAAATCGAACCC 1479
DB 555 AACTTTGTCTCTCCAGAAATTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTC 496
QY 1480 CA 1481
DB 495 AA 494

RESULT 40
BX409035/c
LOCUS
DEFINITION BX409035 Homo sapiens PLACENTA COT 25-NORMALIZED EST 03-MAY-2004
3-PRIME, mRNA sequence.
ACCESSION BX409035

VERSION BX409035.2 GI:46956493
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 484)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 15, 2003 this sequence version replaced gi:30766926.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0BAD005ZE05_AD00422_l&c=4215.r

FEATURES
source Location/Qualifiers
1..484
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 27.1%; Score 401.4; DB 5; Length 484;
Best Local Similarity 91.9%; Pred. No. 3.1e-107;
Matches 445; Conservative 0; Mismatches 37; Indels 2; Gaps 2;
QY 681 GGAAATACCCATACCTCAACCTCACCTGCTGTAATAATTAGCAATACTATAGACACAC 740
DB 484 GGAAAGAACCCATCCCGTCAAGCTCACCTGCTGTAATAATTAGCAATACTATAGACACAC 425
QY 741 CAGCTCCCAATGATCAGGTGGGTAAACACCTCCACACGAATAGTCTGCTACCCCTCAGG 800
DB 424 CAACTCACATGATCAGGTGGGTAACTCTCCACACAAATAGTCTGCTACCCACAGCA 365
QY 801 AATATTTTTTGTCTGTGGTACCTCAGCCTATCATTTGTTGAATGGCTCTTCAGAACTCTAT 860
DB 364 AATATTTTTTGTCTGTGGTACCTCAGCCTATCGTTGTTGAATGGCTCTTCAGAACTCTAT 305
QY 861 GTGCTTCTCTCATTTCTTAGTGGCCCTATGACCTACACTGACACAGATTATACAA 920
DB 304 GTGCTTCTCTCATTTCTTAGTGGCCCTATGACCTACACTGACACAGATTATACAA 245
QY 921 TCATGTCTGTACCTAAGCCCCACAAACAAAGAGTACCCTTCTTCTTTTGTATCAGAGC 980
DB 244 TTATGTCATATCTAAGCCCGGCAACAAGAGTACCCTTCTTCTTTTGTATAGGAGC 185
QY 981 AGGAGTCTAGGAGCAGACTAGGTACTGGCAATTGGCAGTATCAACCTCTACTCAGTTCTA 1040
DB 184 AGGAGTCTAGGAGCAGACTAGGTACTGGCAATTGGCAGTATCAACCTCTACTCAGTTCTA 125
QY 1041 CTACAACTATCTCAAGAAATAATGGTGCATGGGAACAGGTCACTGACT-CCCTGGTCA 1099
DB 124 CTACAACTATCTCAAGAACTAAATGGGACATGGGAACGGGTGCCGACTCCCTCTGGTCA 65
QY 1100 CCTTGCAGATCAACTTAACTCCCTAGCAG-CAGTAGTCTCTCAAAATCGAAGAGCTTTA 1158

Db 64 CCTTGCAGATCAACTTAACTCCCTAGCAGCCAGTAGTCCCTCAAAATNCGAAGAGCTTA 5
Qy 1159 GACT 1162
Db 4 GACT 1
Search completed: February 21, 2005, 14:51:11
Job time : 4614.5 secs

Result No.	Score	♂ Match	Length	DB	ID	Description	
1	1326	99.8	1329	6	BD136199	Retrovira	
2	1326	99.8	1329	6	AR344389	Sequence	
3	1326	99.8	1329	6	AX001030	Sequence	
C	4	1316.4	99.1	163803	9	AC093531	Homo sapi
5	1076.8	81.0	101122	9	AY101590	Pongo pyg	
6	1067	80.3	148711	9	AC133134	Homo sapi	
7	1064.8	80.1	10124	9	AY101531	Pongo pyg	
8	1036.8	78.0	2694	9	AF520477S2	Homo sapi	
9	1036.8	78.0	2694	9	AF520489S2	Homo sapi	
10	1036.8	78.0	2694	9	AF520505S2	Homo sapi	
11	1036.8	78.0	2694	9	AF520529S2	Homo sapi	
12	1035.2	77.9	2694	9	AF520483S2	Homo sapi	
13	1035.2	77.9	2694	9	AF520487S2	Homo sapi	
14	1035.2	77.9	2694	9	AF520493S2	Homo sapi	
15	1035.2	77.9	2694	9	AF520495S2	Homo sapi	
16	1035.2	77.9	2694	9	AF520497S2	Homo sapi	
17	1035.2	77.9	2694	9	AF520503S2	Homo sapi	
18	1035.2	77.9	2694	9	AF520509S2	Homo sapi	
19	1035.2	77.9	2694	9	AF520513S2	Homo sapi	

Source	1. .1329 /organism="unidentified" /mol_type="genomic DNA" /db_xref="taxon:32644"						
Query Match	99.8%; Score 1326; DB 6; Length 1329;						
Best Local Similarity	100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;						
Matches 1329; Conservative	0; Mismatches 0; Indels 0; Gaps 0;						
ORIGIN							
Qy	1	TCAAAATCGAAGAGCTTTAGAGCTTCTAAACCGCCAAAGAGGGGAACTGTTTATTTTT	60				
Db	1	TCAAAATCGAAGAGCTTTAGAGCTTCTAAACCGCCAAAGAGGGGAACTGTTTATTTTT	60				
Qy	61	AGGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAAGA	120				
Db	61	AGGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAAGA	120				
Qy	121	AATTTGAGATCGAATATATGTPAGAGCAGAGACCTTCAAAACACTGCACCTGGGGCCT	180				
Db	121	AATTTGAGATCGAATATATGTPAGAGCAGAGACCTTCAAAACACTGCACCTGGGGCCT	180				
Qy	181	CCTCAGCCAATGGATGCGCTTGAAGCTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT	240				
Db	181	CCTCAGCCAATGGATGCGCTTGAAGCTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT	240				
Qy	241	TTTACTCTCTTTGGACCCCTGATCTTCAACTTCTTGAAGTTTGTCTCTCCAGAAAT	300				
Db	241	TTTACTCTCTTTGGACCCCTGATCTTCAACTTCTTGAAGTTTGTCTCTCCAGAAAT	300				
Qy	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAGTCCATGACTAA	360				
Db	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAGTCCATGACTAA	360				
Qy	361	AATCTACGTGGACCCCTGGACCGGCTGTAGACTGTCTGATGATTAATGACATTGA	420				
Db	361	AATCTACGTGGACCCCTGGACCGGCTGTAGACTGTCTGATGATTAATGACATTGA	420				
Qy	421	AGTCACCCCTCCGAGGAATCTCAACTGTCACAAACCCCTACTACACTCCAAATTCAGTAGG	480				
Db	421	AGTCACCCCTCCGAGGAATCTCAACTGTCACAAACCCCTACTACACTCCAAATTCAGTAGG	480				
Qy	481	AAGCAGTTAGACAGTTGTACGCCAACCTTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540				
Db	481	AAGCAGTTAGACAGTTGTACGCCAACCTTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540				
Qy	541	GGGTGACTCTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAGCCT	600				
Db	541	GGGTGACTCTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAGCCT	600				
Qy	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	660				
Db	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	660				
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720				
Db	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720				
Qy	721	AATCATCTATTGCTCAGAGACACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780				
Db	721	AATCATCTATTGCTCAGAGACACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780				
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGTTTT	840				
Db	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGTTTT	840				
Qy	841	CACCTCTATTTCACCTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTTTATGG	900				
Db	841	CACCTCTATTTCACCTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTTTATGG	900				
Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCTGCTGTTTGGCCAGGTCAGACCCGCT	960				
Db	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCTGCTGTTTGGCCAGGTCAGACCCGCT	960				

Qy	961	GCTGACTTCACATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTGATCCACGCGAGGT	1020				
Db	961	GCTGACTTCACATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTGATCCAGCGAGGT	1020				
Qy	1021	ACCCATTGCCACTCCCGATCAGGCTAAAGGCTTGGCCATTGTCTCTGATGGCTAAGTGCC	1080				
Db	1021	ACCCATTGCCACTCCCGATCAGGCTAAAGGCTTGGCCATTGTCTCTGATGGCTAAGTGCC	1080				
Qy	1081	TGGGTTGTCTCTAATAGAACTGAAACACTGGTCACTGGGTTCCATGGTTCTCTTCCATGAC	1140				
Db	1081	TGGGTTGTCTCTAATAGAACTGAAACACTGGTCACTGGGTTCCATGGTTCTCTTCCATGAC	1140				
Qy	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCTTTGGTA	1200				
Db	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCTTTGGTA	1200				
Qy	1201	TCTGTAGGCGCAAGAACCCCAAGTCAAGAAAGTGAGGCTTGGCCACCATTTGGGAAGTGG	1260				
Db	1201	TCTGTAGGCGCAAGAACCCCAAGTCAAGAAAGTGAGGCTTGGCCACCATTTGGGAAGTGG	1260				
Qy	1261	CCCAGTGCATTTTGGTAGCGGCCACCAACCATCTTGGGAGCTGTGGGACCAAGGATCCC	1320				
Db	1261	CCCAGTGCATTTTGGTAGCGGCCACCAACCATCTTGGGAGCTGTGGGACCAAGGATCCC	1320				
Qy	1321	CCAGTAACA 1329					
Db	1321	CCAGTAACA 1329					

RESULT 2

AR344389

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Sequence 108 from patent US 6582703.

AR344389.1

GI:33740330

Unknown.

Unclassified.

1 (bases 1 to 1329)

Perron, H., Beseme, F., Bedin, F., Paranhos-Baccala, G., Komurian-Pradel, F., Jolivet-Reynaud, C. and Mandrand, B. Isolated nucleotide sequences associated with multiple sclerosis or rheumatoid arthritis and a process of detecting

Patent: US 6582703-A 108 24-JUN-2003;

Location/Qualifiers

1. .1329

/organism="unknown"

/mol_type="mRNA"

Query Match

Best Local Similarity

Matches 1329; Conservative

99.8%; Score 1326; DB 6; Length 1329;

100.0%; Pred. No. 0;

0; Mismatches 0; Indels 0; Gaps 0;

1

1

61

61

121

121

181

181

241

241

301

301

361

361

421

421

481

481

541

541

601

601

661

661

721

721

781

781

841

841

901

901

961

961

Db 241 TTTACTCTCTTGGACCCCTGTATCTTCAACTCTCTGTTAAGTTGTCTCTCCAGAT 300
301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTGATCTGATGTTAATGACATTGA 420
361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTGATCTGATGTTAATGACATTGA 420
421 AGTCACCCCTCCCGAGGAATCTCACTGCGACACCCCTTACTACACTCCAAATTCAGTAGG 480
421 AGTCACCCCTCCCGAGGAATCTCACTGCGACACCCCTTACTACACTCCAAATTCAGTAGG 480
481 AAGCAGTTAGAGCAGTTGTCAGCAACCTCCCAACAGTACTGTTGGTTTTCTGTTTGA 540
481 AAGCAGTTAGAGCAGTTGTCAGCAACCTCCCAACAGTACTGTTGGTTTTCTGTTTGA 540
541 GGTGGACTGAGAGCAGCAGTACTGATGTTTCTAGGCTGACTAAGAAATCCNAGCCT 600
541 GGTGGACTGAGAGCAGCAGTACTGATGTTTCTAGGCTGACTAAGAAATCCNAGCCT 600
601 ANCTGGGAAGGTGACCGCATCTCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
601 ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
661 ACCAATCAGAGCTCACTAAAATGCTAATCAAGCAATGGGGCTTGCAACTTAGCTCACACCG 720
661 ACCAATCAGAGCTCACTAAAATGCTAATCAAGCAATGGGGCTTGCAACTTAGCTCACACCG 720
721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATCTGATGGAGCTCTGTTTT 840
781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATCTGATGGAGCTCTGTTTT 840
841 CACTATTTTCACTATTAATCATGCACTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 900
841 CACTATTTTCACTATTAATCATGCACTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 900
901 CTCACCTGAGCTTTTGTGCGCATCCACACTGCTGTTTGGCACCCTGCTCAGACCCGCT 960
901 CTCACCTGAGCTTTTGTGCGCATCCACACTGCTGTTTGGCACCCTGCTCAGACCCGCT 960
961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCAATGCTGCTGCTGCTGCTGCTGCTGCT 1020
961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCAATGCTGCTGCTGCTGCTGCTGCTGCT 1020
1021 ACCCATGTCATCCCGATCAGGCTAAGGCTTGGCAATGTTCTGCTGCTGCTGCTGCTGCTGCT 1080
1021 ACCCATGTCATCCCGATCAGGCTAAGGCTTGGCAATGTTCTGCTGCTGCTGCTGCTGCTGCT 1080
1081 TGGGTTGCTCTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1081 TGGGTTGCTCTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1141 CCAAGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTTCAATTCCTTGGTA 1200
1141 CCAAGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTTCAATTCCTTGGTA 1200
1201 TCTGTAGGCAAGAACCCAGGTCAGAGAAAGTGGCTTGGCAATTTGGGAAGTGG 1260
1201 TCTGTAGGCAAGAACCCAGGTCAGAGAAAGTGGCTTGGCAATTTGGGAAGTGG 1260
1261 CCCACTGCAATTTGGTAGGGCCCAACCATCTTTGGGAGCTGTTGGGCAAGGATCCC 1320
1261 CCCACTGCAATTTGGTAGGGCCCAACCATCTTTGGGAGCTGTTGGGCAAGGATCCC 1320
1321 CCAAGTAACA 1329

1321 CCAAGTAACA 1329
RESULT 3
AX001030
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match
Best Local Similarity 99.8%; Score 1326; DB 6; Length 1329;
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGGAACTGTTTATTTT 60
DB 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCCAAGAGGGGAACTGTTTATTTT 60
QY 61 AGGGGAAGAAATCTGTTAGTATGTTAATCAATCTGGAATCATCTAGAGAAAGTTAAAGA 120
DB 61 AGGGGAAGAAATCTGTTAGTATGTTAATCAATCTGGAATCATCTAGAGAAAGTTAAAGA 120
QY 121 AATTTGAGATCCAAATATATAGTAGAGCAGAGACCTTCAAAACACTGCACTTGGGCGCT 180
DB 121 AATTTGAGATCCAAATATATAGTAGAGCAGAGACCTTCAAAACACTGCACTTGGGCGCT 180
QY 181 CTTAGCCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CTTAGCCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTTCTTAAAGTTGTCTCTTCCAGAA 300
DB 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTTCTTAAAGTTGTCTCTTCCAGAA 300
QY 301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
DB 301 TGAAGCTGTAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGATGCTCACTAA 360
QY 361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTGATCTGATGTTAATGACATTGA 420
DB 361 AATCTACCTGGACCCCTGGACCGGCTGCTAGATCTGATCTGATGTTAATGACATTGA 420
QY 421 AGTCACCCCTCCCGAGGAATCTCACTGCGACACCCCTTACTACACTCCAAATTCAGTAGG 480
DB 421 AGTCACCCCTCCCGAGGAATCTCACTGCGACACCCCTTACTACACTCCAAATTCAGTAGG 480
QY 481 AAGCAGTTAGAGCAGTTGTCAGCAACCTCCCAACAGTACTGTTGGTTTTCTGTTTGA 540
DB 481 AAGCAGTTAGAGCAGTTGTCAGCAACCTCCCAACAGTACTGTTGGTTTTCTGTTTGA 540
QY 541 GGTGGACTGAGAGCAGCAGTACTGATGTTTCTAGGCTGACTAAGAAATCCNAGCCT 600
DB 541 GGTGGACTGAGAGCAGCAGTACTGATGTTTCTAGGCTGACTAAGAAATCCNAGCCT 600
QY 601 ANCTGGGAAGGTGACCGCATCTCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
DB 601 ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660

601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCGG	660
661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC	720
661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC	720
721	AATCATCTATTGCTGAGAGACAGCGGAGGAGGACAGGATGGGATATAAATCAGGCA	780
721	AATCATCTATTGCTGAGAGACAGCGGAGGAGGACAGGATGGGATATAAATCAGGCA	780
781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTT	840
781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTT	840
841	CACCTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTATGG	900
841	CACCTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTTATGG	900
901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGTCACAGCCGCT	960
901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGTCACAGCCGCT	960
961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACCTGTGCTCCTGATCCAGCGAGT	1020
961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACCTGTGCTCCTGATCCAGCGAGT	1020
1021	ACCCATTGCCATCCCGATCAGGCTAAAGGCTTGCCATTTGCTGCAATGGCTAAAGTCC	1080
1021	ACCCATTGCCATCCCGATCAGGCTAAAGGCTTGCCATTTGCTGCAATGGCTAAAGTCC	1080
1081	TGGGTTTGTCTTAATAGAACTCAACACCTGCTGCTGGTTCATGTTCTTCCATGAC	1140
1081	TGGGTTTGTCTTAATAGAACTCAACACCTGCTGCTGGTTCATGTTCTTCCATGAC	1140
1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA	1200
1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA	1200
1201	TCGTGAGGCCAAGAACCCAGGTCAGAAAGTGAAGCTTCCACCATTTGGGAAGTGG	1260
1201	TCGTGAGGCCAAGAACCCAGGTCAGAAAGTGAAGCTTCCACCATTTGGGAAGTGG	1260
1261	CCCACTGCCATTTTGTAGCGGCCACCACTTGGAGCTGTGGGAGCAAGGATCCC	1320
1261	CCCACTGCCATTTTGTAGCGGCCACCACTTGGAGCTGTGGGAGCAAGGATCCC	1320
1321	CCAGTAACA 1329	
1321	CCAGTAACA 1329	
RESULT 4		
AC093531/c		
LOCUS		
DEFINITION		
AC093531		
AC093531.2		
GI:16945981		
HTG.		
SOURCE		
ORGANISM		
Homo sapiens (human)		
Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
1 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
Direct Submission		
Unpublished		
REFERENCE		
DOE Joint Genome Institute.		
Direct Submission		
Submitted (31-AUG-2001)		
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
3 (bases 1 to 163803)		
DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE		
JOURNAL		
AUTHORS		
DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE		

```

Db      116884 AATCATCTATTGCTGAGAGCACAGCGGAGGACAAGGATTTGGATATAAAACCCAGGCA 116825
QY      781  TTCAAGCAGCAACAGCAACCCCTTTGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 840
Db      116824 TTCAAGCAGCAACAGCAACCCCTTTGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 116765
QY      841  CACTCTATTTCACCTATTAAATCATGCAACTGCACTCTCTGCTCCGTTGTTTTATGG 900
Db      116764 CACTCTATTTCACCTATTAAATCTTGCACTGCACTCTCTGCTCCGTTGTTTTATGG 116705
QY      901  CTCAAGCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGGCACCGTCACAGACCGCT 960
Db      116704 CTCAAGCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGGCACCGTCACAGACCGCT 116645
QY      961  GCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCACCTGCTGCTCTGATCCAGCAGCT 1020
Db      116644 GCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCACCTGCTGCTCTGATCCAGCAGCT 116585
QY      1021 ACCATTGCACTCCCGATCAGGCTAAAGCTTGCATTTGCTTCTGATGGCTTAAGTGC 1080
Db      116584 ACCATTGCACTCCCGATCAGGCTAAAGCTTGCATTTGCTTCTGATGGCTTAAGTGC 116525
QY      1081 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTCTGCTGCTTCCATGCTTCTTCCATGAC 1140
Db      116524 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTCTGCTGCTTCCATGCTTCTTCCATGAC 116465
QY      1141 CCACGGCTTCTAATAGAGCTAATACACTACCGCATGCGCCCAAGATTCATCTCTTGGTA 1200
Db      116464 CCACGGCTTCTAATAGAGCTAATACACTACCGCATGCGCCCAAGATTCATCTCTTGGTA 116405
QY      1201 TCTGTGAGCCAGAACCCAGCTCAGAGAGTGGCTTGCACCATTTGGGAAGTGG 1260
Db      116404 TCTGTGAGCCAGAACCCAGCTCAGAGAGTGGCTTGCACCATTTGGGAAGTGG 116345
QY      1261 CCACCTGCCATTTGTTAGCGGCCACCACTCTTTGGGAGCTGTGGGACGAGGATCCC 1320
Db      116344 CCACCTGCCATTTGTTAGCGGCCACCACTCTTTGGGAGCTGTGGGACGAGGATCCC 116285
QY      1321 CCAGTAACA 1329
Db      116284 CCAGTAACA 116276

RESULT 5
AY101590
LOCUS
DEFINITION
Pongo pygmaeus isolate 1 endogenous retrovirus ERV-W, ERVWE1 locus,
allele A, complete sequence.
ACCESSION
AY101590
VERSION
AY101590.1 GI:37544415
KEYWORDS
SOURCE
Pongo pygmaeus (orangutan)
ORGANISM
Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE
1 (bases 1 to 10122)
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G.,
Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
JOURNAL
PUBMED
2 (bases 1 to 10122)
Mallet, F., Bouton, O. and Oriol, G.
Direct Submission
AUTHORS
Submitted (06-MAY-2002) Retrovirology Department, UMR 2142
JOURNAL
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
1. .10122
Location/Qualifiers
/organism="Pongo pygmaeus"
/mol_type="genomic DNA"
/isolate="1"
FEATURES
source

```

```

/isolation_source="PBM of orangutan 1"
/db_xref="taxon:9600"
/cell_type="PBM"
/notes="isolated by PCR; ERVWE1 locus, allele A
endogenous_virus: ERV-W"
1. .780
Join(247. .811,2872. .9669)
/notes="putative mRNA transcript 1; gag-pol pseudogene"
Join(247. .811,7440. .9669)
/notes="envelope glycoprotein"
/notes="putative mRNA transcript 2"
Join(247. .811,9104. .9669)
/notes="putative mRNA transcript 3; no coding region"
7668. .9284
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68169.1"
/db_xref="GI:37544416"
/translation="MALPYHIFLFTLLPFTLTAPPCCRCMTSSSPYQEFPLRMHHP
GNIDAPYSRFSKGTPTFTAHMPKNCVNSATLQVHSTSPYKGLDLSKLHETLRT
TVCTWYFTHTGSDGGVQDEAREKHVKEISQLTQVHSTSPYKGLDLSKLHETLRT
HRLVSLFNITLUTGLHVSQAQNPNCWMLPLAFRYSIPVPEQNNFSTETNTSV
YRLGSLSESMFLSLVPMPTIYEQDLNYSVSKPRNRPVLPFPVMAAGVGLGALG
TGIGITTSQFYKLSOBLGMDMERVADSLVTLQDLSIAAVLQNRRLDLDLFLP
RGCTCLFLGEECCYVNGSGIVTEKVEIRDRIRRAEELRNIGPMGLFSQWMPWLP
FLGPLAAIILLLLFGPCIFNLAVNPFVSSRIEAVKLQMEPKMQSKTIYHPLDWPASP
RSDVNDIKGTPEEISTAQPLLRPNAGSS"
9343. .10122

LTR
ORIGIN
Query Match      81.0%; Score 1076.8; DB 9; Length 10122;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1186; Conservative 0; Mismatches 130; Indels 13; Gaps 2;
QY      1  TCAAAATCGAAGAGCTTTAGACTTGCTAACCCGCAAAAGAGGGGAAACCTGTTTATTTT 60
Db      8807 TCAAAATCGAAGAGCTTTAGACTTGCTAACCCGCAAAAGAGGGGAAACCTGTTTATTTT 8866
QY      61  AGGGGAAGAAATGCTGTATTGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
Db      8867 AGGGGAAGAAATGCTGTATTGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 8926
QY      121 AATTTGAGATCGAATATAATGTAGACGAGAGGACCTTCAAAACACCTGCACCTGGGCGCT 180
Db      8927 AATTTGAGATCGAATATAATGTAGACGAGAGGACCTTCAAAACACCTGCACCTGGGCGCT 8986
QY      181 CCTCAGCCAAATGATGCCCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 240
Db      8987 CTTACGCAATGATGCCCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 9046
QY      241 TTTACTCTCTTTGGACCTGTATCTCAACTCTCTTGTGTTAAGTTGTCTCTTCCAGAA 300
Db      9047 GCTACTCTCTTTGGACCTGTATCTCAACTCTCTTGTGTTAAGTTGTCTCTTCCAGAA 9106
QY      301 TGAAGCTGTAAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 360
Db      9107 CGAAGCTGTAAAGCTACAATAAGTTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 9154
QY      361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTGTTAATGACATTTGA 420
Db      9155 GATCTACCAACAGACCCCTGGACTGGCTGTCTGCTACCCACGATCTGATGTTAATGACATCAA 9214
QY      421 AGTCACCCCTCCGAGGAAATCTCAACTGTCACAAACCCCTACTACATCTCCAAATTCAGTAGG 480
Db      9215 AGGCACCCCTCCGAGGAAATCTCAACTGTCACAAACCCCTACTACATCTCCAAATTCAGTAGG 9274
QY      481 AAGCAGTTAGAGAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db      9275 AAGCAGTTAGAGAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 9334
QY      541 GGGTGGACTGAGAGACAGGAGTACTAGTGGATTTCTTAGGCTGACTAAGAATCCCAAGCCT 600

```

Db 9335 TGGGGACTGAGAGACAGACTAGCTGATTTCTTAGCCCTACTAAGAAATCCCTAAGCCT 9394
Qy 601 ANCTGGGAAGTACCGCATCATCTTTAAACATGGGCTTGCACACTTAGCTCACACCG 660
Db 9395 AGCTGGGAAGTACCGCATCATCTTTAAACATGGGCTTGCACACTTAGCTCACACCTG 9454
Qy 661 ACCAATCAGAGAGCTCATAAATGCTAATCAGGCAAAAAGGAGGTAAAGCAATAGCC 720
Db 9455 ACCAATCAGAGAGCTCATAAATGCTAATCAGGCAAAAAGGAGGTAAAGCAATAGCC 9514
Qy 721 AATCATCTATTGCTTGCAGACACGCGGAAGCAAGGATTTGGGATATAAATCAGGCA 780
Db 9515 AATCATCTATTGCTTGCAGACACGCGGAAGCAAGGATTTGGGATATAAATCAGGCA 9574
Qy 781 TTCAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGATGAGGAGCTGTGTTT 840
Db 9575 TTCGAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGATGAGGAGCTGTGTTT 9634
Qy 841 CACTCTATTGCTCTATTAAATCATGCAACAGGCTCTTCCCATTTGATGATGAGGAGCTGTGTTT 900
Db 9635 CATGCTATTGCTCTATTAAATCATGCAACAGGCTCTTCCCATTTGATGATGAGGAGCTGTGTTT 9694
Qy 901 CTCAGAGCTGAGCTTTTGGTTCGCTATCCACCTGCTGTTTGGCCACCGTCACAGACCCGCT 960
Db 9695 CTTGAGCTGAGCTTTTGGTTCGCTATCCACCTGCTGTTTGGCCACCGTCACAGACCCGCT 9754
Qy 961 GCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCAGGAGGT 1020
Db 9755 ACTGACTCCCATCCCTCGGATTCCTGCAGGCTGTCCACTGTGCTCTGATCAGGAGGT 9814
Qy 1021 ACCCATTCGCTCCCGATCAGGCTAAAGGCTTGCATTTGCTGTCATGCTTAAGTGCC 1080
Db 9815 GCCCATTTGCGCTCCCAATCAGGCTAAAGGCTTGCATTTGCTGTCATGCTTAAGTGTC 9874
Qy 1081 TGGGTTTGTCTTAATAGAACTGAACA CTGGTCACTGGGTTCATGGTTCCTCTTCATGAC 1140
Db 9875 CAGGTTGCTCTTAATGAGCTGAACA CTAGTCACTGGGTTCACCGTTCCTCTCTGTCAG 9934
Qy 1141 CCAGGCTTCTAATAGAGTATAACA CTACCGGATGCGCCGAGATTCATTCCTTGTA 1200
Db 9935 CCACAGCTTCTAATAGAGTATAACA CTACCGGATGCGCCGAGATTCATTCCTTGTA 9993
Qy 1201 TCTGTAGGCCAAGAACCCAGCTCAGAGAGTGGCTTGCCACCACTTTGGGAAGTGG 1260
Db 9994 TCATGAGGCCAAGAACCCAGCTCAGAGAGTGGCTTGCCACCACTTTGGGAAGTGG 10053
Qy 1261 CCCACTGCCATTTGGTAGCGGCCACCACTTCTGGAGGCTGTGGAGCAAGGATCCC 1320
Db 10054 CCGGCCACCATCTTGAAGCAGCTCACCACCATCTTGGAGGCTGTGTAGCAAGGACCCC 10113
Qy 1321 CCAGTAACA 1329
Db 10114 CAAGTAACA 10122

RESULT 6
AC133134
LOCUS AC133134 148711 bp DNA linear PRI 28-NOV-2002
DEFINITION Homo sapiens chromosome 5 clone RP11-120B7, complete sequence.
ACCESSION AC133134
VERSION AC133134.2 GI:25815299
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 148711)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 148711)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission
Submitted (07-SEP-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 148711)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
Submitted (28-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Nov 28, 2002 this sequence version replaced gi:22758291.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.7.
FEATURES
source
1. 148711
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-120B7"
ORIGIN
Query Match 80.3%; Score 1067; DB 9; Length 148711;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 1175; Conservative 0; Mismatches 143; Indels 4; Gaps 2;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTTGCTTAACCGCCAAAGAGGGGAACCTGTTTATTTT 60
Db 99272 TCAAAATCGAAGAGCTTTAGACTTTGCTTAACCGCCAAAGAGGGGAATCTGTTTATTTT 99331
Qy 61 AGGGGAAGAAATGCTTTAGTATGTTTAAATCAATCTGGAATCAATCTGAGAAGTTAAAGA 120
Db 99332 AAAGAAGAAATGCTTTAGTATGTTTAAATCAATCAGAAATCGTCTGAGAAGTTAAAGA 99391
Qy 121 AATTGAGATCGAATATAATGTAGAGCAGAGAGCTTCAAAACACTGCAACCTGGGGCT 180
Db 99392 AATTGGAATCAATACAACTAGAGCAAAAGAGCTTCAAAACGTTGGGACACTGAGGCT 99451
Qy 181 CCTCAGCCAAATGAGTCCCTGAGCTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 99452 CCTCAGCCAAATGAGTCCCTGAGTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 99511
Qy 241 TTTACTCTCTTTTGAGACCTGTATCTTCAACTTCTTGTTAAGTTGTCTCTTCAGAAAT 300
Db 99512 GCTACTCATTTTAGACCTCTATCTTTAACTCTCTTAAAGTTGTCTCTTCAGAAAT 99571
Qy 301 TGAAGCTGAAGCTTCAAAATAGTCTTCAAAATGGAACCCAGATGCCAGTCCATGACTAA 360
Db 99572 CGAAGCTGTAATACTTCAAAATCATTTCTTCAAAATGGAGGCCAGATGCCAGCTAA 99631
Qy 361 AATCTACCGTGAGCCCTGGAGCCCTGCTAGACTATGCTCTGATGTTTAATGACATTGA 420
Db 99632 GATCTACCGTGAGCCCTGGAGCCCTGCTAGACTATGCTCTGATGTTTAATGACATTGA 99691
Qy 421 AGTCACCCCTCCGAGGAAATCTCAACTGSCAAACCCCTACTTACACTCCAAATTCAGTAG 480
Db 99692 AGGCACCCCTCCGAGGAAATTTCAACTGSCAAACCCCTACTTACACCCCAATTCAGCAG 99751
Qy 481 AAGCAGTTAGAGCAGTTGTTCAGCCCAACCTCCCAACAGTACTTGGGTTTCCCTGTTGAGA 540
Db 99752 AAGCAGTTAGAGTGTGTCTCGGCCAACCTCCCAACAGTACTTGGGTTTCCCTGTTGAGA 99811
Qy 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 99812 GGGGTACTGAGAGCAGGACTAGCTGGATTTTCTTAGGCCGACTAAGAAATCCCTAAGCCT 99871
Qy 601 ANCTGGGAAGGTGACCCGATCCATCTTTAAACATGGGCTTGCACACTTAGCTCACACCG 660
Db 99872 AGCTGGGAAGGTGACCCGATCCATCTTTAAACATGGGCTTGCACACTTAGCTCACACCTG 99931
Qy 661 ACCAATCAGAGAGCTCATAAATGCTAATCAGGCAAAAAGGAGGTAAAGCAATAGCC 720


```
QY 481 AAGCAGTTAGACAGCTGTGTGAGCAACCTCCCAACAGTACTTGGTGTCTCTGTTGAGA 540
|
|
|
Db 9275 AAGCAGTTAGACAGCTGTGTGAGCAACCTCCCAACAGTACTTGGTGTCTCTGTTGAGA 9334
|
|
|
QY 541 GGGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCNNAAGCCT 600
|
|
|
Db 9335 TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCCTAAGCCT 9394
|
|
|
QY 601 ANCTGGGAAGGTGACCGGATCCATCTTTAAACATGGGGCTTGCAACTAGCTCACACCG 660
|
|
|
Db 9395 AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTAGCTCACACCTG 9454
|
|
|
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
|
|
|
Db 9455 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 9514
|
|
|
QY 721 AATCATCTATTGCTCTGAGAGCACAGCGGGAAGGACAAGGATTTGGATATATAAATCAGGCA 780
|
|
|
Db 9515 AATCATCTATTGCTCTGAGAGCACAGCGGGAAGGACAATGATCGGGATATATAAATCAGGCA 9574
|
|
|
QY 781 TTCAAGCCGACGACAGCAACCCCTTTGGGT--CCCTCCCAATTTGATGGGAGCTCTGTT 838
|
|
|
Db 9575 TTCAAGCCGACGACAGCAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTT 9634
|
|
|
QY 839 TTCACTCTATTGCTCTGAGAGCACAGCGGGAAGGACAAGGATTTGGATATATAAATCAGGCA 898
|
|
|
Db 9635 TTCACTCTATTGCTCTGAGAGCACAGCGGGAAGGACAATGATCGGGATATATAAATCAGGCA 9694
|
|
|
QY 899 GGCTCAAGCTGAGCTTTGTCGCCATCCACACTGCTGTTTGGCCACCGTACAGACCGG 958
|
|
|
Db 9695 GGCTCAAGCTGAGCTTTGTCGCCATCCACACTGCTGTTTGGCCACCGTACAGACCGG 9754
|
|
|
QY 959 CTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAG 1018
|
|
|
Db 9755 CCACCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAG 9814
|
|
|
QY 1019 GTACCCATTGCCACTCCGATCAGCTAAAGCTTTGCCATTTCTGCTGATGGCTAAGTG 1078
|
|
|
Db 9815 GTGCCATTGCCCTCCGATCAGCTAAAGCTTTGCCATTTCTGCTGATGGCTAAGTG 9874
|
|
|
QY 1079 CTGCGGTTGCTCTAATGAACTGAACACTGCTGCTAAGCTTTGCCATTTCTCTTCCATG 1138
|
|
|
Db 9875 TCCAGGTTGCTCTAATGAGCTGAACACTGCTGCTAAGCTTTGCCATTTCTCTTCCATG 9934
|
|
|
QY 1139 ACCCAGGCTTCTAATAGCTTATACACTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1198
|
|
|
Db 9935 ACCCAGGCTTCTAATAGCTTATACACTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9993
|
|
|
QY 1199 TATCTGTGAGGCAAGAACCCAGGTCAGAGAAAGTGGGCTTGGCCATTTGGGAAGT 1258
|
|
|
Db 9994 AATCCATGAGGCAAGAACCCAGGTCAGAGAAAGTGGGCTTGGCCATTTGGGAAGT 10053
|
|
|
QY 1259 GGCACCATGCTGCTTTGTTAGCGGCGCCACCACTCTTGGGAGCTGTGGGAGCAAGATC 1318
|
|
|
Db 10054 AGCCGCGCCACCATTTGGAAGCAGCTCAGCAGCTTTGGGAGCTCTGTGAGCAAGGACC 10113
|
|
|
QY 1319 CCCCAGTAACA 1329
|
|
|
Db 10114 CCCCAGTAACA 10124
|
|
|
```

```
RESULT 8
AF52047782
LOCUS AF52047782 2694 bp DNA linear PRI 11-FEB-2004
DEFINITION Homo sapiens individual 132 allele A, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520478
VERSION AF520478.1 GI:33410924
KEYWORDS 2 of 2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 2694)
Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G.,
Bonnaud B., Lucotte G., Duret L., and Mandrand B.
The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
2 (bases 1 to 2694)
Mallet F., Bouton O., and Oriol G.
Direct Submission
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-BioMérieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 132 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="Caucasian"
endogenous_virus: HERV-W"
misc_feature 43..57
/note="splice acceptor site"
CDS 284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA017561.1"
/db_xref="GI:33410925"
/translaton="MALPYHFLFTVLLPSFLTAPPCCRTSSSPYQEFELWQRP
GNIDAPSYLSKGTPTTATHMPNRYHSATLCHMANTHYTGKMLNPSCPGLGV
TVCTYPTQGMDSGGVQDQAEKHEVKSQILTRVHTGSSPPKGLDLSKLHETLR
LTVLVSFNLTTLGLHEVSAQNPNCWICLPNRPYVSIPIVQMNFFSTVTSV
LVGLVSNLEITHTNLTKVFSNTVTTSQIRVWTPPTQIVCLPSGIFVQGTSA
YRLNGSESMCPLSFLVPMWTIVTEODLYNVIKPNKRPVLPFVIGAGVLGALG
TGGITTSITQFYKLSQELNGMDERVADSLVTLDOLNLAVALNRRALDLTAE
RGGCLFLPGECCYKNSGIVSTBKVRIDRIORRAEHLRNTGFWGLLSQWMPHLP
FLGPLAAIIILLGLPCIFNLNLFVSRIBAEKLVQMEPKMQSKTKIYRRFLDRPAS
RSDVNDIKGTPPEISAAQPLLRNSAGSS"
misc_feature 1705..1720
/note="splice acceptor site"
LTR 1959..2694
ORIGIN
Query Match 78.0%; Score 1036.8; DB 9; Length 2694;
Best Local Similarity 90.2%; Pred. No. 9.7e-305;
Matches 1134; Conservative 0; Mismatches 110; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTTATTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGAACTGTTATTTT 1482
QY 61 AGGGGAAGATGCTGTAGTAGTATGTTAATCAATCGGAATCATCTAGAGAAGTTAAAGA 120
Db 1483 AGGGGAAGATGCTGTAGTAGTATGTTAATCAATCGGAATCATCTAGAGAAGTTAAAGA 1542
QY 121 AATTGAGATCGAATATAATGTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 1543 AATTCGAGATCGAATACAACTAGAGCAGAGGAGCTTCGAAACACTGCAGACCTGGGGCT 1602
QY 181 CCTCAGCAATGATGACCTGACCTCCCTCTCTTAGGACCTCTAGCAGCTATAATT 240
Db 1603 CCTCAGCAATGATGACCTGACCTCCCTCTCTTAGGACCTCTAGCAGCTATAATT 1662
QY 241 TTTACTCTCTTTGGACCTCTATCTTCAACTTCTTAAAGTTTGTCTCTTCCAGAT 300
Db 1663 GCTACTCTCTTTGGACCTCTATCTTCAACTTCTTAAAGTTTGTCTCTTCCAGAT 1722
```


301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA 360
 1723 CGAAGCTGTAAACTA-----CAATGAGCCCAAGATGCGAGTCCAAAGCTAA 1770
 361 AATACCGTGGACCCCTGGACCGCCCTGCTAGACTATGCTCTGATGTTAATGACATGGA 420
 1771 GATCTACCGACAGCCCTGGACCGCCCTGCTAGCCACCGATCTGATGTTAATGACATCAA 1830
 421 AGTACCCCTCCGAGGAAATCTCACTGACAAACCCCTACTACACTCCAAATTCAGTAGG 480
 1831 AGGACCCCTCTCTGAGGAAATCTCACTGACAAACCCCTACTACCGCCCAATTCAGCAGG 1890
 481 AAGCAGTTAGACAGTTGTGACCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
 1891 AAGCAGTTAGAGGGTCTGTCGGCAACCTCCCAACAGCAGTACTGTTTCTCTGTTGAGA 1950
 541 GGGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAGGCT 600
 1951 TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT 2010
 601 ANCTGGGAGGTGACCGGATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCG 660
 2011 AGCTGGGAGGTGACACATCCACCTTTAAACACGGGGCTTGCACCTTAGCTCACACCTG 2070
 661 ACCAATCAGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
 2071 ACCAATCAGAGAGCTCACTAAATATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
 721 AATCATCTATTGCTGAGAGCAAGCGGGAAGCAAGGATTTGGGATATAAATCTCAGGCA 780
 2131 AATCATCTATTGCTGAGAGCAAGCAGGAGGACAAATGATCGGGATATAAACCACCAAGTC 2190
 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGATGGAGCTCTGTTTT 840
 2191 TTGAGCCGCAACAGCAACCCCTTTGGGTCCTCCCTCCCTTTGATGGAGCTCTGTTTT 2250
 841 CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGCTTTTATGG 900
 2251 CATGCTATTTCATCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTACGG 2310
 901 CTAAGCTGAGCTTTTGTTCGCAATCACCACTGCTGTTTGGCAACCGTCACAGACCCGCT 960
 2311 CTTGAGCTGAGCTTTGCTCGCATCAACCACTGCTGTTTGGCCACCGCAGACCCGCT 2370
 961 GCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCACTGTGCTCTCTGATCCAGCAGGT 1020
 2371 GCTGACTCCATCCCTCTGGATCATGACGGGTGTCCGCTGTGCTCTGATCCAGCAGAGC 2430
 1021 ACCATTGGCAGTCCCGATCAGGCTAAAGGCTTGCATTTGTTCTGCTGCTTAAGTGC 1080
 2431 ACCATTGGCGTCCCAATCGGCTAAAGGCTTGCATTTGTTCTGCTGCTTAAGTGC 2490
 1081 TGGGTTTGTCTTAATAGAACTGAACCTGGTCACTGGGTTCCATGTTCTTCTCCATGAC 1140
 2491 TGGGTTTCTTAATGAGTGAACCTAGTCACTGGGTTCCATGTTCTTCTCTGAC 2550
 1141 CACGCTTCTTAATAGAGTATTAACACTACCGCATGGCCCAAGATTCCTTCTGTTA 1200
 2551 CCACAGCTTCTTAATAGAGTATTAACACTACCGCATGGCCCAAGGTTTCCATCTT-GAA 2609
 1201 TCTGTGAGGCAAGAACCCAGGTCAGAGAAGTGGGCTTGGCACCATTGGGAAG 1257
 2610 TCCATAGGCAAGAACCCAGGTCAGAGAACACGAGGGCTTGGCACCATTGGGAG 2666

RESULT 9
 AF520489S2 2694 bp DNA linear PRI 11-FEB-2004
 LOCUS Homo sapiens individual 22 allele A, envelope glycoprotein gene,
 DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
 ACCESSION AF520490
 VERSION AF520490.1 GI:33410948
 KEYWORDS

2 of 2
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
 Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
 TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
 in hominoid placental physiology
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
 PUBMED 14757826
 REFERENCE 2 (bases 1 to 2694)
 AUTHORS Mallet,F., Bouton,O. and Oriol,G.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée
 d'Italie, Lyon 69364 cedex 07, France
 FEATURES
 source
 1..2694
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="individual 22 allele A"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-q22"
 /sex="male"
 /cell_type="PBMC's"
 /note="Caucasian"
 endogenous_virus: HERV-W
 43..57
 /note="splice acceptor site"
 284..1900
 /note="syncytin"
 /codon_start=1
 /product="envelope glycoprotein"
 /protein_id="AAQ17567.1"
 /db_xref="GI:33410949"
 /translation="MALPYHFLFTVLLPSFTLTAPPCCRCMTSSSPYOEFLWRMORP
 GNIDAPSVRSLSKGTPTTAHTMPNCHYSATLQMHANTHYMTGKMLNPSGPGIGV
 TVCTYFTGMSDGGVQDQAREKHVEVISQLTRVHGTSSPYGDLDSLKHETLRT
 HTRLVSLFNTTLGLHEVSAQNPNCWICLPLNFRPYVSIIPVEQNNFSTHEINTSV
 LVGLVSLNLEITHTSNLTVCVFSTNTYTNQSCIRWPTPTQICLPGSIFVCGTSA
 YRLNGSSSECSFLSPVPMPTIYTEQDLYNYVISKPRKRVPLIPFVIGAGVGLGALG
 TGIGITSTOFYKLSOELGMDERVADSLVTLQDLNLAALVQNRALDLTLAE
 RGCTCLFGEBCCYVNGSGIVTEKVEIRDIRORAEELRTGPKGLLSOMPMPLP
 FLGPLAAILLLLPFCIFNLLVNFVSSRIENAVKLQMEPKMQSKTKIIRPLDRPAS
 RSDVNDIKGTPPEISAAQPLLRPNASGSS"
 1705..1720
 /note="splice acceptor site"
 1959..2694

Query Match 78.0%; Score 1036.8; DB 9; Length 2694;
 Best Local Similarity 90.2%; Pred. No. 9.7e-305;
 Matches 1134; Conservative 0; Mismatches 110; Indels 13; Gaps 2;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTCTACCCCAAGAGGGGGAACCTGTTATTTT 60
 DB 1423 TCAAAATCGAAGAGCTTTAGACTTCTACCCCTGAAAGAGGGGGAACCTGTTATTTT 1482
 QY 61 AGGGGAAGAGTCTGTTAGTATGTTAATCAATCTGGAATCATTACTGAGAAAAGTTAAAGA 120
 DB 1483 AGGGGAAGAGTCTGTTATGTTAATCAATCCGGAATCGTCACTGAGAAAAGTTAAAGA 1542
 QY 121 AATTTGAGATCGAATATAATGTAGAGCAGAGACCTTCAAAACACTGCACCTCGGGCCCT 180
 DB 1543 AATTCGAGATCGAATACAACGCTAGAGCAGAGAGCTTCGAAACACTGGACCTCGGGCCCT 1602
 QY 181 CCTCAGCCATGATGATCCCTCGGACTCTCCCTCTTAGGACCTCTAGCAGCTATATATT 240
 DB 1603 CCTCAGCCATGATGATGATCCCTCTTAGGACCTCTTAGGACCTCTAGCAGCTATATATT 1662

complete cds, and 3' long terminal repeat, complete sequence.	
AF520506	AF520506.1 GI:33410980
VERSION	
KEYWORDS	
SEGMENT	
SOURCE	
ORGANISM	
2 of 2	
Homo sapiens (human)	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 2694)	
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.	
The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology	
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	
14757826	
2 (bases 1 to 2694)	
Mallet, F., Bouton, O. and Oriol, G.	
Direct Submission	
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France	
Location/Qualifiers	
1. .2694	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/isolate="individual 37 allele A"	
/db_xref="taxon:9606"	
/chromosome="7"	
/map="7q21-q22"	
/sex="male"	
/cell_type="PBMC's"	
/notes="endogenous_virus: HERV-W"	
43..57	
/notes="splice acceptor site"	
284..1900	
/notes="syncytin"	
/codon_start=1	
/product="envelope glycoprotein"	
/protein_id="AA017575.1"	
/db_xref="GI:33410981"	
/translation="MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPVQSFRLMRQRP GNIDAPSYRLSKGTPTFTATHMPKNCYSAILCMHANTHYTGKMINPSCPGGLGV TVCWYPTQTMGSDGGVQDOAREKHKEVISQLTRVHGSSTYTGKGLDLSKLHETLT HTLVLNLFNTLTGLHEVSAQNFTNCWICLPLNFRPYVSIPEQWMNFSTEINTTSLVPLVNLBIHTTSLNLTCKVFNSTYITNSQICIRWYTPQIVCLPSGIFPVCGTSA YRCLNGSSSMCELSFLVPPMTIYTEODLYNVTISKPNKRVPLIPVIGAGVLDLTAETGGITTSQFYFKLSQELNGDMERVADSLVTLQDOLNLAALVQLNRALDILLTAE RGGTCLFLGSECCYYVNSGIVTEKVEIRDRIQORAEELRNTGSPWGLSQWMPWILP FLGPLAAIILLFLFGPCIFNLLVNFVSSRIEAVKIQMEPKMQSKTKYRRPLDRPASP RSDVNDIKGTPTPEEISAAQPLLRPNSSAGSS"	
1705..1720	
/notes="splice acceptor site"	
1959..2694	
misc_feature	
LTR	
ORIGIN	
Query Match 78.0%; Score 1036.8; DB 9; Length 2694;	
Best Local Similarity 90.2%; Pred. No. 9.7e-305;	
Matches 1134; Conservative 0; Mismatches 110; Indels 13; Gaps 2;	
QY	1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAGAGGGGGAACCTGTTTATTTT 60
Db	1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGGAACCTGTTTATTTT 1482
QY	61 AGGGGAAGAAATGCTGTGTAGTATGTTAATCAATCTGGAATCATATTCTAGAAAGTTAAAGA 120
Db	1483 AGGGGAAGAAATGCTGTGTATATGTTAATCAATCCGAATCGTCACTGAGAAGTTAAAGA 1542
QY	121 AATTGAGATCGAATATTAATGTAGACGACGACCTTCAAAACACTGCACCTGGGGCT 180
Db	1543 AATTGAGATCGAATATCAACGCTAGACGACGAGAGCTTCGAAACACTGCACCTGGGGCT 1602
QY	181 CCTCAGCCCAATGATGCCCTCGGACTCTCCCTCTCTTAGGACTCTCTAGCAGCTATAATATT 240
QY	241 TTTACTCTCTTTTGGACCCCTGTATCTTCAACTTCTTGTAAAGTTTGTCTTCCAGAAAT 300
Db	1663 GCTACTCTCTTTTGGACCCCTGTATCTTAACTCTTGTAACTTGTCTCTTCCAGAAAT 1722
QY	301 TGAAGCTGTAAAGCTACAAATAGTCTTCAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db	1723 CGAAGCTGTAAAGCTA-----CAATGGAGCCCAAGATGAGTCCCAAGACTAA 1770
QY	361 AATCTACCGTGGACCCCTGGACCGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db	1771 GATCTACCGAGACCCCTGGACCGCTGTAGCCACGATCTGATGTTAATGACATCAA 1830
QY	421 AGTCACCCCTCCGAGGAATCTCACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
Db	1831 AGGCACCCCTCCTGAGGAATCTCAGCTGCACAACCTCTACTAGCCCCCAATTCAGCAGG 1890
QY	481 AAGCAGTTAGACGAGTCTGACGCAACCTCCCAACAGTACTTGGTGTCTCTGTGAGA 540
Db	1891 AAGCAGTTAGACGAGTCTGCGCCAACTCCCAACAGCACTTAGTGTCTCTGTGAGA 1950
QY	541 GGGTGGACTGAGACAGCACTAGCTGGATTTCTCTAGGCTGACTAAGAAATCCCNAAAGCCT 600
Db	1951 TGGGGACTGAGACAGCACTAGCTGGATTTCTCTAGGCTGACTAAGAAATCCCTAAGCCT 2010
QY	601 ANCTGGGAAGTGACCGATCATCTTTAAACATGGGCTTGCACCTAGCTACACCG 660
Db	2011 AGCTGGGAAGTGACCAATCACCTTTAAACACGGGGCTTGCACCTAGCTACACCTG 2070
QY	661 ACCAATCAGAGAGCTCACTAAATCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db	2071 ACCAATCAGAGAGCTCACTAAATGCTAATTAGCAAAAACAGGAGGTAAAGAAATAGCC 2130
QY	721 AATCATCTATTGCTTGAGACACAGCGGAAGGACAAGATTGGATATTAACCTCAGCA 780
Db	2131 AATCATCTATTGCTTGAGACACAGCAGGAGGACAATGATCGGGATATAAACCAAGTC 2190
QY	781 TTCAGGCAGCAACAGCAACCCCTTGGGTCCCTCCCATTTGATGGAGCTGTGTTT 840
Db	2191 TTCAGGCAGCAACAGCAACCCCTTGGGTCCCTCCCTTGTGATGGAGCTGTGTTT 2250
QY	841 CACTCTATTCTACTTATTAATCATGCAACTGCACTCTCTGTGTCCTGTTTTTATGG 900
Db	2251 CATGCTATTCTACTTATTAATCTTGCAACTGCACTCTCTGTGTCCTGTTCTTAGCG 2310
QY	901 CTCAGCTGAGCTTTTGTGGCATCCACCACTGCTGTTGSCACCGTCACAGACCCGCT 960
Db	2311 CTTGAGCTGAGCTTTGCTGCGCATCCCACTGCTGTTGCGCGCACCGCAGACCCGCC 2370
QY	961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCAGGT 1020
Db	2371 GCTGACTTCCATCCCTCTGGATCATGCAAGGTGTCCGCTGTCTCTGATCCAGCAGGC 2430
QY	1021 ACCCATTTGCCACTCCCGATCAGGCTAAAGGTTGCAATGTTCTTCGATGGCTAAGTGC 1080
Db	2431 ACCCATTTGCCCTCCCAATCGGCTAAAGGCTTGCATTTGTTCTTCGATGGCTAAGTGC 2490
QY	1081 TGGGTTTGCCTAATAGAACTGACACTGCTCAGTGGTTCCTGCTCTCTTCCATGAC 1140
Db	2491 TGGGTTTGCCTAATTTGAGCTGAACACTAGTCACTGGGTTCCATGTTCTCTCTGTAC 2550
QY	1141 CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA 1200
Db	2551 CCACAGTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGGTTCCATTCCTT-GAA 2609
QY	1201 TCTGTAGGCCAAGAACCCAGGTCAGAGAANGTAGGCTTGCACCATTTGGGAG 1257
Db	2610 TCCATAAGGCCAAGAACCCAGGTCAGAGAACACGAGGCTTGCACCATCTTGGGAG 2666
RESULT 10	
AF520505S2	
LOCUS	
DEFINITION Homo sapiens individual 37 allele A, envelope glycoprotein gene,	
AF520505S2 2694 bp DNA linear PRI 11-FEB-2004	

```
Db 1603 CTTACGCCAATGATGCCCTCGATCTCTCCCTCTTAGGACTCTAGAGCTATATATT 1662
QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTCTTGTAGTTGCTCTTCCAGAAAT 300
Db 1663 GCTACTCTCTTTGGACCTGTATCTTAACTCTCTTGTAACTTGTCTCTTCCAGAAAT 1722
QY 301 TGAAGCTGTAAAGCTACAATAGTCTTCAAAATGGAACCCAGATGCCAGTCCATCACTAA 360
Db 1723 CGAAGCTGTAAAGCTA-----CAAAATGGAGCCCAAGATGCCAGTCCAAAGCTAA 1770
QY 361 AATCTACCTGTGACCCCTCGACCGGCTCTAGACTATCTCTGTATCTTGAATGAATGA 420
Db 1771 GATCTACCCAGACCCCTGGACCGGCTCTAGCCACGATCTGATTTAATGACATCAA 1830
QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCAACACCCCTACTACACTTCCAAATCAGTAGG 480
Db 1831 AGGCACCCCTCTGAGGAAATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGAGAGTTGTGAGCAACCTCTCCCAACAGTACTTGGGTTTCTCTGTGAGA 540
Db 1891 AAGCAGTTAGAGAGTTGTGAGCAACCTCTCCCAACAGTACTTGGGTTTCTCTGTGAGA 1950
QY 541 GGTGTGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 1951 TGGGGAGCTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGCAACTTACTAGCTCAACCCG 660
Db 2011 AGCTGGGAAGGTGACCATCTCACTTAAACAGCGGGCTTGCAACTTACTAGCTCAACCCG 2070
QY 661 ACCAATCAGAGAGCTCACTAAATGCTATTCAGGCAAAAACAGGAGTTAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGCTCACTAAATGCTATTCAGGCAAAAACAGGAGTTAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTCTGAGAGCAGCGGGAAGGACAAGGATTTGGGATATAAACTCAGGCA 780
Db 2131 AATCATCTATTGCTCTGAGAGCAGCGGGAAGGACAAGGATTTGGGATATAAACTCAGGCA 2190
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCACTGTATGGAGGCTCTCTTTT 840
Db 2191 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCACTGTATGGAGGCTCTCTTTT 2250
QY 841 CACTCTATTTCATCTTAAATCATGCACTGCACTCTCTGCTCGGTCTTTTATGG 900
Db 2251 CATGCTATTTCATCTTAAATCTTGCACTGCACTCTCTGCTCGGTCTTTTATGG 960
QY 901 CTCAAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGTCAACAGCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGTCAACAGCCGCT 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGGT 1020
Db 2371 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGGT 2430
QY 1021 ACCATTGCACTCCCGATCAGGCTTAAAGCTTGCATGTTTCTCTGATGCTTAAAGTCC 1080
Db 2431 ACCATTGCACTCCCGATCAGGCTTAAAGCTTGCATGTTTCTCTGATGCTTAAAGTCC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 2491 TGGGTTTGTCTTAATAGAACTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2550
QY 1141 CCACGGCTTCTAATAGAGCTATAAAGCTCAACCGCATGCGCCCAAGATTCCTATCTTGGTA 1200
Db 2551 CCACGGCTTCTAATAGAGCTATAAAGCTCAACCGCATGCGCCCAAGATTCCTATCTTGGTA 2609
QY 1201 TCTGTGAGGCAAGAACCCAGGTCAGAGAAAGTGGGCTTGGCAACATTTGGGAG 1257
Db 2610 TCCATAGGCAAGAACCCAGGTCAGAGAAAGTGGGCTTGGCAACATTTGGGAG 2666
```

RESULT 11

```
AF520529S2 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 78 allele A, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520530
VERSION AF520530.1 GI:33411028
KEYWORDS
SEGMENT
SOURCE
ORGANISM
2 of 2
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVW1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE
2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL
FEATURES
Location/Qualifiers
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 78 allele A"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="Caucasian"
endogenous_virus: HERV-W"
43..57
/note="splice acceptor site"
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17587.1"
/db_xref="GI:33411029"
/translation="MALPYHIFLVLPSPFTLTAPPCCRMVTSSTSPYQSFELRMORP
GNIDAPYSRSLSGTPTFTATHMPNRCVTHSATLCMHANTHWTKMINPSCPGGLGV
TYCVTFYFTGMSDGGVQDQAREKHVEVISQLTRVHGTSSPYKGLDLSKLHETLRT
HTRVLSLNTLTGLHEVSAQNPTNCWICLPLNFRPYVSIPIVPEQNNFSTEINTSV
LVGPLVNLITHTSLNLTCKFSNTYTTNSQICIRWVTPPTQICLPSGIFVCGTSA
VCLNGSSSCFLSLVPPMTIYTEQDLYNVYISKPRNRPVILPFVIGAGVIGALG
TGIGITTSQFYKLSOELNGDMERVADSLTLODOLNSLAUVI.ONRRALDOLLAE
RGCTCLFLGEBECYYVNSGIVTEKVEIRDRIRKAEELRNTGPNGLLSQWMPHILP
FLGPLAAIILLLFGPCIFNLNLFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP
RSDVNDIKGTPPEISAAQPLLRPSAGSS"
1705..1720
/misc_feature
/note="splice acceptor site"
1959..2694
ORIGIN
Query Match 78.0%; Score 1036.8; DB 9; Length 2694;
Best Local Similarity 90.2%; Pred. No. 9,7e-305;
Matches 1134; Conservative 0; Mismatches 110; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAAACCTGTTATTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCGTAAGAGAGGGGAAACCTGTTATTTT 1482
QY 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCATTACTGAGAAAGTAAAGA 120
Db 1483 AGGGGAAGATGCTGTTAGTATGTTAATCAATCCGGAATCGTCACTGAGNAAGTAAAGA 1542
QY 121 AATTTGAGATCGAATATAATATGAGCAGAGGACCTTCAAAACACTGCACCTGGGCGCT 180
```

Db 1543 AATTCGAGATCGATACACGTAGACAGAGAGCTTCGAACACTGACCTGGGCGCT 1602
QY 181 CCTCAGCAATGGATGGCCCTGGACCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 240
Db 1603 CCTCAGCAATGGATGGCCCTGGATCTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 1662
QY 241 TTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTCTGTTAAAGTTTGTCTCTCCAGAA 300
Db 1663 GCTACTCTCTTTGGACCCCTGATCTTTAACTCTCTGTTAACTTTGCTCTTCCAGAA 1722
QY 301 TGAAGCTGTAAGCTTACAAATAGTTCTTTCAAAATGGAACCCAGATGCACTCCATGACTAA 360
Db 1723 CGAAGCTGTAAGCTA-----CAATGGAGCCCAAGATGCACTCCAAGACTAA 1770
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTGA 420
Db 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACTCAA 1830
QY 421 AGTCAACCCCTCCGAGGAAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGG 480
Db 1831 AGGCACCCCTCTGAGGAAATCTAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGACAGTGTCTGACCAACCTCCCAACAGTACTTGGGTTTTCTGTGTGAGA 540
Db 1891 AAGCAGTTAGAGCGGTGCTGGCCAACTCTCCCAACAGCACTTAGGTTTTCTGTGTGAGA 1950
QY 541 GGTGGACTGAGACAGACGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCNAAGCCT 600
Db 1951 TGGGGACTGAGACAGACGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCCTAAGCCT 2010
QY 601 ANCTGGGAAGGTACCGCATCATCTTTAAACATGGGCTTGCACCTTAGCTCACACCG 660
Db 2011 AGCTGGGAAGGTGACCAATCCCTTTAAACACGGGCTTGCACCTTAGCTCACACCTG 2070
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGAGACAGCGGAGGAGCAAGGATGGATATAACTCAGGCA 780
Db 2131 AATCATCTATTGCTGAGAGACAGCAGGAGGAGCAATGATCGGATATAAACCAAGTC 2190
QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTGATGATGGAGCTGTGTTT 840
Db 2191 TTCGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTGATGATGGAGCTGTGTTT 2250
QY 841 CACTCTATTTCACCTATTAAATCATCAACTGCACTCTTCTGCTGCCGTGTTTTATGG 900
Db 2251 CATGCTATTTCACCTATTAAATCTTGCAACTGCACTCTTCTGCTGCCATGTTCTTACGG 2310
QY 901 CTCAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCT 960
Db 2311 CTTGAGCTGAGCTTTTCGCTCGCATCCACCACTGCTGTTTGGCCCAACCGCAGACCGCC 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGCTCACTGCTCTCTGATCCAGCGAGGT 1020
Db 2371 GCTGACTCCCATCCCTCTGGATCATGCAAGGTGTCGCTGCTGCTCTGATCCAGCGAGGC 2430
QY 1021 ACCCATTTGCCATCCCGATCAGGCTAAAGGCTTGCCATTTGTTCTGCAATGCTAAGTGCC 1080
Db 2431 ACCCATTTGCCCTCCCAATCGGCTAAAGGCTTGCCATTTGTTCTGCAATGCTAAGTGCC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTTCCATGAC 1140
Db 2491 TGGGTTTCACTTAAATTTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGAC 2550
QY 1141 CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGTGTA 1200
Db 2551 CCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTT-GAA 2609
QY 1201 TCTGTGAGGCCAAGAACCCAGGTGAGGAANGTGAGGCTTGCCACCAATTTGGGGAAG 1257

Db 2610 TCCATAAGGCCAAGAACCCAGGCTCAGAGAACAGAGGCTTGCCACCATCTTGGGAG 2666

RESULT 12
AF520483S2 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 147 allele B, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
AF520484
AF520484.1 GI:33410936
KEYWORDS
SEGMENT
SOURCE 2 of 2
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,G., Duret,L. and Mandrand,B.
Bonnaud,B., Lucotte,G., Prudhomme,S., Cheynet,V., Oriol,G.,
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
JOURNAL 14757826
PUBMED 2 (bases 1 to 2694)
REFERENCE Mallet,F., Bouton,G. and Oriol,G.
AUTHORS Direct Submission
TITLE Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
JOURNAL CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69384 cedex 07, France
LOCATION/Qualifiers
FEATURES
source
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 147 allele B"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
misc_feature 43..57
/note="splice acceptor site"
CDS 284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17584.1"
/db_xref="GI:33410937"
/translation="MALPYHIFLFTVLLPSFTLTAPPPCRMTSSSPYQBELWRMQR
P
GNIDAPSYRLSKGTPFTTAHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLV
TICWYFTQTMGSDGGVQDAQREKHVKEVISQLTRVHGTSYSPYKGLDLSKLHETLR
HTRLVSNLTTLGLHVSANPTNCWICLPNRPVYIPVPEQMNFSFEINTVSV
LVGPLVSNLEIHTHSLTCKFSNTTNTNSQCIHWTPPTQIIVCLPSGIPFVCGTSA
YRLGSESMCFSLFVLPMTIYTEQDLYSVISKPRKRVPLIPVIGAGVLGALG
TGIGITTSQFYKJQSQELNGMVADELVTLDQQLNSLAALVQLNRRALDLTAE
RGSTFLPGCEYKYNQSGIVTEKVEIKDRIQRRABELRNTGFWLLSQMMPWILP
PLGPLAAIIILLFLGPCIIFNLVNFVSSRIEAVKLQMEPKMQSKTKIYRRDLRPA
PS
RSDVNDIKGTTPPEISAAQPLLRNSAGSS"
misc_feature 1705..1720
/note="splice acceptor site"
LTR 1959..2694
ORIGIN
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAGAGGGGAACTGTTATTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGTGAAGAGGGGAACTGTTATTTT 1482
QY 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCAATCTGGAAGTAAAGA 120

Db 1423 TCATAATCGAAGAGCTTTAGACTTGCTAAACGCTGAAGAGGGGAAACCTGTTATTTT 1482

Qy 61 AGGGAAGAATGCTGTTAGTAGTTAATCAATCTGGAATCATATTACTGAGAAGTTAAAGA 120

Db 1483 AGGGAAGAATGCTGTTATATTATTAATCAATCTGGAATCTGCTACTGAGAAGTTAAAGA 1542

Qy 121 AATTGAGATCGATATATATGCTAGAGCAGAGACCTTCAAAACACCTGCACCCCTGGGGCT 180

Db 1543 AATTGAGATCGAATACAACTGAGAGCAGAGAGCTTCGAAACACCTGCACCCCTGGGGCT 1602

Qy 181 CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240

Db 1603 CCTCAGCAATGGATGCGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 1662

Qy 241 TTTACTCTCTTTGAGCCCTGATCTTCAACTCTGTTTAAAGTTTGTCTCTCCAGAAAT 300

Db 1663 GCTACTCTCTTTGAGCCCTGATCTTAAACCTCTCTGTTAACTTGTCTCTCCAGAAAT 1722

Qy 301 TGAAGCTGAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCACTCCATGACTAA 360

Db 1723 CGAAGCTGAAGCTA-----CAATGGAGCCCAAGATGCACTCCAGACTAA 1770

Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420

Db 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830

Qy 421 AGTCACCCCTCCGAGGAAATCTCACTGACACACCCCTACTACACTCAATTCTAGTAGG 480

Db 1831 AGGCACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 1890

Qy 481 AAGCAGTTAGAGCAGTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTGAGA 540

Db 1891 AAGCAGTTAGAGCGGTGCTGGCCAACTCTCCCAACAGACTTAGTTTCTCTGTGAGA 1950

Qy 541 GGGTGGACTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTAAGAAATCCNNAAGCT 600

Db 1951 TGGGGGACTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTAAGAAATCCNNAAGCT 2010

Qy 601 ANCTGGGAAGTGACCGCATCATCTTTAAACATGGGCTTGCACTTAGCTACACCCG 660

Db 2011 AGCTGGGAAGTGACCAATCCACTTTAAACACGGGGCTTGCACTTAGCTACACCTG 2070

Qy 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAAAGGAGGTAAAGCAATAGCC 720

Db 2071 ACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAAAAGGAGGTAAAGCAATAGCC 2130

Qy 721 AATCATCTATTCCCTGAGAGCAGCGGGAAGCAAGATTTGGATATATAACTCAGCA 780

Db 2131 AATCATCTATTCCCTGAGAGCAGCAGGAGGCAATGATCGGGATATAAACCAAGTC 2190

Qy 781 TTCAAGCCAGCAACAGCAACCCCTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840

Db 2191 TTCAAGCCAGCAACAGCAACCCCTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 2250

Qy 841 CACTCTATTCTACTTATTAAATCATGCAACTGCACTCTCTTGGTCCGTTTATGG 900

Db 2251 CATGCTATTCTACTTATTAAATCTTGCAACTGCACTCTCTTGGTCCATGTTTCTACGG 2310

Qy 901 CTCAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTTGGCCACGCTCAGACCCGCT 960

Db 2311 CTTGAGCTGAGCTTTTCGCTCGCATCCACCACTGCTGTTTGGCCACGCTCAGACCCGCT 2370

Qy 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTCCACTGCTCTCTGATCCAGCGAGT 1020

Db 2371 GCTGACTTCCATCCCTCTCGATCATGAGGGTCTCCGCTGCTCTGATCCAGCGAGG 2430

Qy 1021 ACCCATTTGCCACTCCCGATCAGGCTAAAGGCTTGCCATTGTTCTCGCATGGCTAAAGTCC 1080

Db 2431 ACCCATTTGCCCTCCCAATTTGGCTTAAAGGCTTGCCATTGTTCTCGCATGGCTAAAGTCC 2490

Qy 1081 TGGGTTTGTCTTAATAGAACTGAACCTGCTCAGTGGGTCCATGGTCTCTTCTCCATGAC 1140

Db 2491 TGGGTTTGTCTTAATAGAACTGAACCTAGTACCTGGGTCCATGGTCTCTTCTCTGTGAC 2550

Qy 1141 CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGTA 1200

Db 2551 CCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGTTCCATTCCTT-GAA 2609

Qy 1201 TCTGTGAGGCCAAGAACCCCGAGGTACAGAAAGTGAGGCTTGCCACCATTTGGGAAG 1257

Db 2610 TCCATAAGGCCAAGAACCCCGAGGTACAGAAACACGAGGCTTGCCACCATCTTGGGAG 2666

RESULT 14

AF520493S2 2694 bp DNA linear PRI 11-FEB-2004

LOCUS Homo sapiens individual 24 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.

DEFINITION AF520494 AF520494.1 GI:33410956

ACCESSION AF520494

VERSION AF520494.1

KEYWORDS 2 of 2

SEGMENT

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2694)

AUTHORS Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.

TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

PUBMED 14757826

REFERENCE 2 (bases 1 to 2694)

AUTHORS Mallet, F., Bouton, O. and Oriol, G.

TITLE Direct Submission

JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France

FEATURES

source

1. .2694

/organism="Homo sapiens"

/mol_type="genomic DNA"

/isolate="individual 24 allele A"

/db_xref="taxon:9606"

/map="7q21-q22"

/sex="female"

/cell_type="PBMC's"

/note="Asian

endogenous_virus: HERV-W

43. .57

/note="splice acceptor site"

284. .1900

/codon_start=1

/product="envelope glycoprotein"

/protein_id="AAQ17569.1"

/db_xref="GI:33410957"

/translation="MALPYHIFLFTVLLPSFTLTAPPRCMTSSSPYQBELWRNQR GNIDPVSRLSKGTPFTAHTHMPNCSYHATLCMHANTHTYTKMINPSCPGGLGV TVCKYFTQTMDSGGVQDOAREKHVKVVISOLTRVHGTSSPYKGLDLKSLHETLR HTRLVSNLEITHSTNLTCTVRFSPNTTNTNSQCIHWTPPTQIIVCLPSGIIFVCGTSA LVGLVNSGSMCFSLFVPPMTIYEQDLYSVISKPNKRPVILPFIYIGAVLGLGAL YRCLNGSTQFYKLSQELNMDMERVADSLVTIQDQLNSLAAVVLQNRALDLITAE TGIQGITTTQFYKLSQELNMDMERVADSLVTIQDQLNSLAAVVLQNRALDLITAE RGGTCLFLGECYVYVQSGIIVTEKVKIIRDIQRABELRNTGFWGLLSQWMPWILP FGFLAAIILLLFGPCITNLLVNFVSRIRIEAVKLQMBKMQSKTKIYRRPLDRPASP RSDVNDIKGTPPEISAAQPLLRPNAGSS"

1705. .1720

/note="splice acceptor site"

1959. .2694

misc_feature

CDS

Query Match 77.9%; Score 1035.2; DB 9; Length 2694;

Best Local Similarity 90.1%; Pred. No. 3e-304;

ORIGIN

Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;	
QY 1 TCAAAATCGAAGAGCTTTAGACTGCTAAACCCGCAAAAGAGGGGAACTGTTTATTTT 60	
Db 1423 TCAAAATCGAAGAGCTTTAGACTGCTAAACCCGCTGAAAGAGGGGAACTGTTTATTTT 1482	
QY 61 AGGGGAAGAGTCTGTTAGTATGTTAACTCAATCTGAATCATTTACTTGAGAAAGTTAAAGA 120	
Db 1483 AGGGGAAGAGTCTGTTATGTTTAACTCAATCTGAATCTGCTAGTGAAGAGTTAAAGA 1542	
QY 121 AATTTGAGATCGAATATATATATAGAGCAGAGGAGCTTCAAAACACTGCAACCTGGGGCT 180	
Db 1543 AATTCGAGATCGAATACAGCTAGAGCAGAGGAGCTTCGAAACACTGGACCCCTGGGGCT 1602	
QY 181 CCTAGCCAAATGGATGCCCTGGAAGCTCTCCCTTCTTAGGACCTCTAGAGCTATATATT 240	
Db 1603 CCTAGCCAAATGGATGCCCTGGAATCTCCCTTCTTAGGACCTCTAGAGCTATATATT 1662	
QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTGTTAAAGTTTGTCTCTTCCAGAA 300	
Db 1663 GTACTCTCTTTGGACCTGATCTTAACTCTCTTGTGTTAAAGTTTGTCTCTTCCAGAA 1722	
QY 301 TGAAGCTGTAAGACTCAAAATAGTTCTTCAAAATGGAACCCAGATGCACTCCATGACTAA 360	
Db 1723 CGAAGCTGTAAGACTA-----CAATGGAGCCCAAGATGCACTCCAGACTAA 1770	
QY 361 AATCTACCGTGAACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGGA 420	
Db 1771 GATCTACCGGAGACCCCTGGACCGGCTGCTAGCCCAAGATGCTGATGTTAATGACATCAA 1830	
QY 421 AGTCACCTCTCCGAGGAATCTCACTGACCAACCCCTACTACACTCCAAATTCAGTAGG 480	
Db 1831 AGGCACCTCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890	
QY 481 AAGCAGCTTAGAGCAGTGTGTCAGCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540	
Db 1891 AAGCAGTTAGAGGCTGCTGGCAACCTCCCAACAGCAGCTTGGTTTTCTGTTGAGA 1950	
QY 541 GGGTGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGTGACTAAGATCCCAAGCT 600	
Db 1951 TGGGGGACTGAGAGACAGACTAGCTGGAATTTCTTAGGTGACTAAGATCCCAAGCT 2010	
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTAAAGATGGGCTTGCACCTAGCTCACACCCG 660	
Db 2011 AGCTGGGAAGGTGACCACTCCACCTTAAACAGGGGCTTGAACCTAGCTCACACCTG 2070	
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 720	
Db 2071 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 2130	
QY 721 AATCATCTATTGCTGAGAGCAGCGGGGAAGCAAGGATGGGATATAAATCTCAGGCA 780	
Db 2131 AATCATCTATTGCTGAGAGCAGCGGGGAAGCAAGGATGGGATATAAATCTCAGGCA 2190	
QY 781 TTAAGCCAGCACAGCAACCCCTTTGGTCCCTCCCTCCATCTGATGGGAGCTCTGTTT 840	
Db 2191 TTAAGCCAGCACAGCAACCCCTTTGGTCCCTCCCTCCCTCTGATGGGAGCTCTGTTT 2250	
QY 841 CACTTATTTCACTCTATTAAATCATGCAACTGCACCTCTCTGCTCGCTGTTTTATGG 900	
Db 2251 CATGCTATTTCACTCTATTAAATCATGCAACTGCACCTCTCTGCTCGCTGTTTTATGG 2310	
QY 901 CTCAGCTAGCTTTGTTTCCGCAATCAACACTGCTGTTTGGCCACCGTCAAGACCCGCT 960	
Db 2311 CTCAGCTAGCTTTGTTTCCGCAATCAACACTGCTGTTTGGCCACCGTCAAGACCCGCT 2370	
QY 961 GCTGACTTCCATCCCTTTGATCCAGAGAGTCTCCTGCTGCTCTGATCCAGCAGGCT 1020	
Db 2371 GCTGACTTCCATCCCTTTGATCCAGAGAGTCTCCTGCTGCTCTGATCCAGCAGGCT 2430	
QY 1021 ACCATTGGCACTCCCGATCAGGCTTAAGGCTTGGCAATGCTTCTGCTGATGCTTAAGTCC 1080	
Db 2431 ACCATTGGCGCTCCCGATCAGGCTTAAGGCTTGGCAATGCTTCTGCTGATGCTTAAGTCC 2490	

QY 1081 TGGGTTTCTCTTAATAGAACTGAACACCTGGTCACTGGGTTCCATGGTCTCTTCCATGAC 1140	
Db 2491 TGGGTTTCTCTTAATAGAACTGAACACCTGGTCACTGGGTTCCATGGTCTCTTCCATGAC 2550	
QY 1141 CCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA 1200	
Db 2551 CCACAGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTT-GAA 2609	
QY 1201 TCTGTGAGCCCAAGAACCCAGGTCAGAGAANGTGCAGGCTTCCACCACTTTGGGAAG 1257	
Db 2610 TCCATAAGGCCCAAGAACCCAGGTCAGAGAACACAGGCTTCCACCACTTTGGGAG 2666	

RESULT 15	
AF520495S2	2694 bp DNA linear PRI 11-FEB-2004
LOCUS	Homo sapiens individual 24 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.
DEFINITION	
ACCESSION	AF520496
VERSION	AF520496.1 GI:33410960
KEYWORDS	
SEGMENT	2 of 2
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 2694)
TITLE	Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B. The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED	14757826
REFERENCE	2 (bases 1 to 2694)
AUTHORS	Mallet,F., Bouton,O. and Oriol,G.
TITLE	Direct Submission
JOURNAL	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France
FEATURES	Location/Qualifiers
source	1..2694
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/isolate="individual 24 allele B"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7q21-q22"
	/sex="female"
	/cell_type="PBMC's"
	/note="Asian endogenous virus: HERV-W"
misc_feature	43..57
	/note="splice acceptor site"
CDS	284..1900
	/note="syncytin"
	/codon_start=1
	/product="envelope glycoprotein"
	/protein_id="AAQ17570.1"
	/db_xref="GI:33410961"
	/translation="MALPYHIEFTVLLPSFTLTAPPCCRCMTSSSPYOEFLWRMQR GNIDAPSVRLSGKSTPTTAHTMPRNCVHSATLCMHANTHYTKMNPSCPGGLV TWYFTTQKSGDGVQDOAREKHVKEVI SOLTRVHGTSSPYKGLDLSKLHETLR TNYVPLVSLTGLHESVQAQNTWCICLPNFRPYVSI PVPEOWNNSTEINTTTSV HRLVSLFTNLTGLHESVQAQNTWCICLPNFRPYVSI PVPEOWNNSTEINTTTSV YRLGSSSMSCFLSFLVPMPTIYEQDLYSVI SKPRNKRVPILDFVIGAGVIGALG TGGITTTSTQFYKLSQBLNGDMERVDLSVLTQDLNSLAALVQLNRRLDLTPAE RGCTCLFGECCYVNOGSI VTEKVEIRDRIOORAEELRNTGPMGLLSQWMPITLP FLGPLAAIILLLELFGPCIFNLAVNFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPAS RSDVNDIKGTPPEEISAAQPLRPNSAGSS"
misc_feature	1705..1720
	/note="splice acceptor site"
LTR	1959..2694

ORIGIN	Query Match Best Local Similarity Matches 1133; Conservative	77.9%; Score 1035.2; DB 9; Length 2694; 90.1%; Pred. No. 3e-304; 0; Mismatches 111; Indels 13; Gaps 2;
Qy	1	TCAAATCGAAGAGCTTTAGACATGCTTAAACGCCCAAGAGGGGACCTGTTATTTT 60
Db	1423	TCAAATCGAAGAGCTTTAGACATGCTTAAACCGCTGAAGAGGGGACCTGTTATTTT 1482
Qy	61	AGGGGAAGAATGCTGTTAGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA 120
Db	1483	AGGGGAAGAATGCTGTTATTTATTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542
Qy	121	AATTTAGATCGAATATAATGTAGACAGAGGACCTTCAAAAACACTGCACCCCTGGGGCCT 180
Db	1543	AATTCGAGATCGAATACAACGTAGACAGAGGAGCTTCGAAAAACACTGGACCCCTGGGGCCT 1602
Qy	181	CCTCAGCCAATGGATGCGCTCGAGACTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT 240
Db	1603	CCTCAGCCAATGGATGCGCTGGATTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT 1662
Qy	241	TTTACTCCTCTTTGGAACCTGTATCTTCAACTCTCTTGTAAAGTTTGTCCTCTCCAGAAT 300
Db	1663	GCTACTCCTCTTTGGAACCTGTATCTTTAAACCTCCTTGTAACTTTGTCCTCTCTCCAGAAT 1722
Qy	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCACAGATGCAGTCCATGACTAA 360
Db	1723	CGAAGCTGTAAACTA-----CAATATGGAGCCCAAGATGCAGTCCAAAGACTAA 1770
Qy	361	AATCTACCGTGGACCCCTCGACCGGCTGCTAGACTATGCTCTGAAGTTTAATGACATTTGA 420
Db	1771	GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830
Qy	421	AGTCACCCCTCCGAGGAATCTCAACTGCACACCCCTACTACACTCCAATTCAGTAGG 480
Db	1831	AGGCACCCCTCCTCAGGAAATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGCAGG 1890
Qy	481	AAGCAGTTTAGACAGTTGTGACGCCAACCTCCCAAACAGTACTTGGGTTTTCTCTGTTGAGA 540
Db	1891	AAGCAGTTAGCGTGTGTCGGCCAACTCCCAACACACTTAGGTTTTCTCTGTTGAGA 1950
Qy	541	GGTGGACTGAGAGACAGGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCCNAGCCT 600
Db	1951	TGGGGGACTGAGACAGGACTAGCTGGAATTTCTCTAGGCTGACTAAGAAATCCCTAAGCCT 2010
Qy	601	ANCTGGGAAGTGACCGCATCTTTAAACATGCGGCTTGGAACTTAGCTCACACCCG 660
Db	2011	AGCTGGGAAGTGACCACTCCACTTTAAACCGGGGCTTGCNACTTAGCTCACACCTG 2070
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAAACAGAGGTTAAGCAATAGCC 720
Db	2071	ACCAATCAGAGAGCTCACTAAATGCTTAATTAGGCAAAAGACAGAGGTTAAGAAATAGCC 2130
Qy	721	AATCATCTATTGCTGAGACAGCGGAGGACAAGGATGCGGATATAAATCTCAGGCA 780
Db	2131	AATCATCTATTGCTGAGACAGAGGAGGACAATGATCGGGAATATAAACCCAAATC 2190
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCACTGTATGCGGAGCTCTGTTTT 840
Db	2191	TTCAAGCCGCAACAGCAACCCCTTTGGGTCCTCCCTCCCTTTGTATGCGGAGCTCTGTTTT 2250
Qy	841	CACCTCATTTTCACCTATTAAATCATGCAACTGCACTCTTCTGGTCTGTTTTTTATGG 900
Db	2251	CATGCTATTTCACCTATTAAATCTTCAACTGCACTCTTCTGGTCTCATGTTCTTTACGG 2310
Qy	901	CTCAAGCTGAGCTTTTGTGTGCCCATCAACACTGCTGTGTGGCAACCGCTCACAGACCCGCT 960
Db	2311	CTTGAGCTGAGCTTTGCTGTGCCCATCAACACTGCTGTGTGGCCACCGGACACCGGCC 2370
Qy	961	GCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCACCTGTGCTCTCGATCCACGAGGT 1020
Db	2371	GCTGACTTCCATCCCTCTGGATCATGAGGGGTGTCGCTGTGCTCTTGATCCACGAGGC 2430


```

RSDVNDIKGTPPEISAAQPLLRPNAGSS"
misc_feature 1705..1720
LTR /note="splice acceptor site"
1959..2694
ORIGIN
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;

QY 1 TCAAAATCGAAGAGCTTTAGACTGTCTAAACCGCAAAAGAGGGGAACTGTGTTATTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTGTCTAAACCGCTGAAAGAGGGGAACTGTGTTATTTT 1482
QY 61 AGGGGAAGATGCTGTATGTTATTAATCAATCTGAATCATTACTGAGAAAGTTAAAGA 120
Db 1483 AGGGGAAGATGCTGTATGTTATTAATCAATCTGAATCATTACTGAGAAAGTTAAAGA 1542
QY 121 AATTTGAGATCGAATATATGAGACGAGGAGCTTTCAAAACACTGCAACCTTGGGGCCT 180
Db 1543 AATTTGAGATCGAATATATGAGACGAGGAGCTTTCAAAACACTGCAACCTTGGGGCCT 1602
QY 181 CCTCAGCCCAATGATGCCCTGGACTCTCCCTTTCTTGGACCTCTAGCAGCTATATATTT 240
Db 1603 CCTCAGCCCAATGATGCCCTGGACTCTCCCTTTCTTGGACCTCTAGCAGCTATATATTT 1662
QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTCTTGTAGTTTGTCTCTTCCAGAAAT 300
Db 1663 GCTACTCTCTTTGGACCTGTATCTTCAACTCTTGTAGTTTGTCTCTTCCAGAAAT 1722
QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA 360
Db 1723 CGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA 1770
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGTATGTTAATGACATGA 420
Db 1771 GATCTACCGGAGACCCCTGGACCGGCTGCTAGCCACGATCTGTATGTTAATGACATCAA 1830
QY 421 AGTCACCCCTCCGAGAGAAATCTCAACTGCAGAACCCCTACTACATCCCAATTCAGTAGG 480
Db 1831 AGGCACCCCTCTTGAGAGAAATCTCAGTGCAGAACCCCTACTACGCCCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGACAGTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db 1891 AAGCAGTTAGACAGTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 1950
QY 541 GGTGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT 600
Db 1951 TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCTTAAACATGAGGGCTTGCACTAGCTCAGACCCG 660
Db 2011 AGCTGGGAAGGTGACCGCATCTTAAACATGAGGGCTTGCACTAGCTCAGACCCG 2070
QY 661 ACCAATCAGAGCTCAGTAAATGCTTAATCAGCAAAACAGAGGTAAGCAATAGCC 720
Db 2071 ACCAATCAGAGCTCAGTAAATGCTTAATGAGCAAGAGAGGTAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGAGCAGACGGGGAAGCAAGGATGGGATATAAACTCAGGCA 780
Db 2131 AATCATCTATTGCTGAGAGCAGACGGGGAAGCAAGGATGGGATATAAACTCAGGCA 2190
QY 781 TTCAAGCAGCAACAGAACCCCTTTGGTCCCTCCCTCCCTTTGATGGGAGCTCTGTTT 840
Db 2191 TTCAAGCAGCAACAGAACCCCTTTGGTCCCTCCCTCCCTTTGATGGGAGCTCTGTTT 2250
QY 841 CACTCTATTTTCACTCTATTAAATCATGCACTCTCTCTGCTCGCTGTTTTTATGG 900
Db 2251 CATGCTATTTTCACTCTATTAAATCATGCACTCTCTCTGCTCGCTGTTTTTATGG 2310
QY 901 CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTACAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTGCTCGCCATCCACCACTGCTGTTTGGCCACCGTACAGACCCGCT 2370

```

```

RESULT 17
AF52050352 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 34 allele B, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520504
VERSION AF520504.1 GI:33410976
KEYWORDS 2 of 2
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
Location/Qualifiers
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 34 allele B"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
misc_feature 43..57
/note="splice acceptor site"
CDS
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17574.1"
/db_xref="GI:33410977"
/translaton="MALPYHIEFLTVLLPSFTLTAPPCCRCMTSSSPYQEFLLWRMORP
GNIDAPSVRLSKGTPTTAAHMPRNCVHSATLCWHANTHYTKMNPSCPGGLGV
TWYVFTQTGMSDGGVQDQAREKHKEVISQLTRVHCTSSPYKGLDLSKHETLRT
HYRLVFNFTLTGLHVSQAQPTNCWICPLNFRPVSIVPVEQWNNFSTEINTTSV
LVGPLVSNLEITHLTCVKFSNTTNTNSQCIRWVTPTQIVCLPSGFFVCGVTSIA
YRLNGSSBSMCFLSFLVPMPTIYTBODLYSYVISKPRNRPVILFPVIGAGVILGALG

```

TGIGGITTSTQFYKLSQELNGDMERVADSLVTLQDOLNSLAAVVLQNRALDLTAE
RGGCLFLGECYVQSGIVTEKVEIRDRIORRAEBELNRTGFWLLSQMWPWILP
PLGFLAAIILLLLFPICIFNLVNFVSRIBAVKLQMEPKMQSKTKIYRRLDRPAS
RSDVNDIKGTPTPEBISAAQPLLRPNAGSS"
1705..1720
/note="splice acceptor site"
1959..2694

misc_feature

LTR

ORIGIN

Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCMAAAGAGGGGAACTGTATATTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCTGAAAAGAGGGGAACTGTATATTTT 1482
QY 61 AGGGAAGAGTCTGTAGTATGTTAAATCAATCTGGAATCAATCTAGTGAAGAAGTTAAAGA 120
Db 1483 AGGGAAGAGTCTGTATGTTTAAATCAATCCGGAATCGTCACTGAGAAGTTAAAGA 1542
QY 121 AATTGAGATCGAATATATATGATAGACGAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 1543 AATTGAGATCGAATATACAAGTAGACGAGAGGAGCTTCGAAACACTGGACCTGGGGCT 1602
QY 181 CCTAGCAATGAGTGCCTGGACTCTCCCTCTTAGGACTCTAGCAGCTATATATTT 240
Db 1603 CCTAGCAATGAGTGCCTGGACTCTCCCTCTTAGGACTCTAGCAGCTATATATTT 1662
QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTCTCTGTTAGTTGTTCTCTTCCAGAA 300
Db 1663 GCTACTCTCTTTGGACCTGATCTTAACTCTCTGTTAACTTGTCTCTTCCAGAA 1722
QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAAGTCCATGATA 360
Db 1723 CGAAGCTGTAAAGCTA-----CAATGAGGCCAAGATGCAAGTCCAAGACTAA 1770
QY 361 AATCTACGTTGACCCCTGGACCGCTGCTAGACTATGCTCTGATGTTATGACATTGA 420
Db 1771 GATCTACGCGACACCCCTGGACCGCTGCTAGCCAGATCGATGTTAATGACATCAA 1830
QY 421 AGTCACTCTCCGAGGAAATCTCAACTGCACAACTCTACTACCTCACTCCAACTCAGTAGG 480
Db 1831 AGGCACTCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGACAGTGTGTCGCCAACCTTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
Db 1891 AAGCAGTTAGACGGGTGTCGCCAACCTTCCCAACAGCAGTACTTGGGTTTTCTGTTGAGA 1950
QY 541 GGGTGACTGAGACGAGTACTGTTGATTTCTAGGCTGACTAAGATCCCNAGCT 600
Db 1951 TGGGGGACTGAGACGAGTACTGTTGATTTCTAGGCTGACTAAGATCCCNAGCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCTTTAAACATGGGCTTGCAACTAGCTCACACCG 660
Db 2011 AGCTGGGAAGGTGACCAATCACTTTAAACAGGGGCTTGCAACTAGCTCACACCTG 2070
QY 661 ACCAATCAGAGAGTCACTAAATGCTAATCAGGCAAAACAGGAGGTTAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGTCACTAAATGCTAATAGGCAAAACAGGAGGTTAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGACGACGCGGAGGACAAGGATTTGGGATATAAATCAGGCA 780
Db 2131 AATCATCTATTGCTGAGACGACGAGGAGGACATGATCGGATATAAATCAGGCA 2190
QY 781 TTCAAGCCAGCAACACCCCTTTGGGTTCCCTCCCATTTGATGGAGCTCTGTTTT 840
Db 2191 TTCAAGCCAGCAACACCCCTTTGGGTTCCCTCCCATTTGATGGAGCTCTGTTTT 2250
QY 841 CACTCTATTCTCTATTAAATCATGCACTGCACTCTCTGGTCCGTTTATATGG 900
Db 2251 CATGCTATTCTCTATTAAATCTTGGCAACTGCACTCTCTGGTCCATGTTTCTTACGG 2310

QY 901 CTAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTTGCCACCGCTCAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTTGCCACCGCTCAGACCCGCT 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGCTCTCTGATCCAGCAGGT 1020
Db 2371 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGCTCTCTGATCCAGCAGGT 2430
QY 1021 ACCATTGCGCACTCCCGATCAGGCTAAAGGCTTGCCATTTGCTGATGCTAAAGTGCC 1080
Db 2431 ACCATTGCGCGCTCCCAATCGGCTAAAGGCTTGCCATTTGCTGATGCTAAAGTGCC 2490
QY 1081 TGGTTTTGTCTTAATAGAACTGAACACTGGTTCATCTGGGTTCCATGTTCTTCCATGAC 1140
Db 2491 TGGTTTTGTCTTAATAGAACTGAACACTGGTTCATCTGGGTTCCATGTTCTTCTGAC 2550
QY 1141 CCAGCGTTCTTAATAGAGCTTAACACTCAGGATGCGCCAGATTCATCTCTTGGTA 1200
Db 2551 CCAGCGTTCTTAATAGAGCTTAACACTCAGGATGCGCCAGATTCATCTCTTGGTA 2609
QY 1201 TCTGTGAGGCCAAGAACCCAGGTCAGGAAGTGGAGCTTGCCACCATTTGGGAAG 1257
Db 2610 TCCATAAGGCCAAGAACCCAGGTCAGGAAGTGGAGCTTGCCACCATTTGGGAAG 2666

RESULT 18

AF52050982

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..2694

/organism="Homo sapiens"

/mol_type="genomic DNA"

/isolate="individual 45 allele A"

/db_xref="taxon:9606"

/chromosome="7"

/map="7q21-q22"

/sex="male"

/cell_type="PBMK's"

/note="endogenous_virus: HERV-W"

43..57

/note="splice acceptor site"

284..1900

/note="syncytin"

/codon_start=1

/product="envelope glycoprotein"

/protein_id="AAQ17577.1"

/db_xref="GI:33410989"

/translation="MALPYHIFLFTVLLPSFTLAPPCCMTSSSPYQEFLLRMQRP
GNIDAPSVRLSKGTPPTTAHTMPRNCVHSATLQMHANTWTKMINPSCPGGLGV
TVCWTFYTIQTGMSDGGVQDQAREKHVKEVISQLTRVHGTSSPYKGLDLSKLHETLRT

```
HTRLVSLFNTLTGLHEVSAQNPTNCWICLPLNFRPVYVSIIPVEOMNFFSTINTTSS
LVGPLVSNLEIHTHTSNLTCVKFSTNTYTNSQIRWVTPPTQIVCLPSGIFPVGCTSA
YRCLNGSSSEMCFLSFLVPPMTIYTEQDLYSVIVISKPNKVPVLPILPFVIGAGVLGALG
TGIGITSTIYQYKLSBELGDMERVADSLVTLQDQNSLAADVVLQNRALDILLTAE
RGCTFLIGBECYVNSGIVTEKVRIRRIORRABELRNTPGWLLSOWMPPILP
FLGPLAAIILLLLFPQCFINLLVNFVSSRIEAVKLQMBPFQMSKTKIYRRPLDRPAS
RSDVNDIKGTPEEBEISAQPLLRPNASGSS"
1705..1720
misc_feature
/note="splice acceptor site"
1959..2694
LTR
ORIGIN
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGCTAAACCCGAAAGAGGGGAAACCTGTTATTTT 60
Db 1423 TCAAAATCGAAGAGCTTTAGACTTGCTAAACCCGCTGAAGAGGGGAAACCTGTTATTTT 1482
QY 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCATTAAGAGAAAGTTAAAGA 120
Db 1483 AGGGGAAGAAATGCTGTATTTAATCAATCGGAATCGTCACTGAGAAAAGTTAAAGA 1542
QY 121 AATTTGAGATCGAATATATGATAGACGAGAGACCTTTCAAAACACTGACCCCTGGGGCT 180
Db 1543 AATTCGAGATCGAATATCAACGATAGACGAGAGGCTTTCGAAACACTGGACCCCTGGGGCT 1602
QY 181 CCTCAGCCCAATGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 1603 CCTCAGCCCAATGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1662
QY 241 TTTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGTAGTTTGTCTCTTCAGAAAT 300
Db 1663 GCTACTCTCTTTGGACCTGTATCTTAACTCTCTTGTAACTTTGTCTCTTCCAGAAAT 1722
QY 301 TGAAGCTGTAAAGTACAAATAGTTCTTCAATGGAACCCGAGATCCAGTCCATGACTAA 360
Db 1723 CGAAGCTGTAAAGTACAAATAGTTCTTCAATGGAACCCGAGATCCAGTCCATGACTAA 1770
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTA 420
Db 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830
QY 421 AGTCACCCCTCCCGAGAAATCTCAACTGACAAACCCCTACTACACTCCAAATTCATGAG 480
Db 1831 AGGCACCCCTCTCTGAGGAAATCTCAGTGCACAAACCTCTACTACGCCCAATTCAGCAG 1890
QY 481 AAGCAGTTGAGCAGTTGTGACGAAACCTCCCAACAGTACTTTGGTCTTCTGTTGAGA 540
Db 1891 AAGCAGTTGAGCGGTGCTGCGGCAACCTCCCAACAGCAGCTTAGTTTCTGTTGAGA 1950
QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT 600
Db 1951 TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGAACCTAGCTAGCTACACCCG 660
Db 2011 AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGAACCTAGCTAGCTACACCCG 2070
QY 661 ACCAATCAGAGCTCAGTAAATGCTTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 720
Db 2071 ACCAATCAGAGCTCAGTAAATGCTTAATGAGCAAAAGCAGAGGTAAGAAATAGCC 2130
QY 721 AATCATCTATTGCTGAGACACAGGGGGAAGGACAAAGGATTTGGATATAAACTCAGGCA 780
Db 2131 AATCATCTATTGCTGAGACACAGGAGGAGGACAAATGATCGGGATATAAAACCAAGTC 2190
QY 781 TTCAGCGCAACAGCAACCCCTTTGGGTCCCTCCCAATGATTTGATGGGAGCTCTGTTTT 840
Db 2191 TTCAGCGCGCAACAGCAACCCCTTTGGGTCCCTCCCTCTTGTATGGGAGCTCTGTTTT 2250
QY 841 CACTCTATTTCATCTATTAAATCATGCACTCTTCTCGTCCGTGTTTTTTTATGG 900
```

```
Db 2251 CATGCTATTTCACCTCTATTAAATCTTGCAACTGCACCTCTCTGGTCCATGTTTCTTAGCG 2310
QY 901 CTCAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTCGCTCGCCATCCACCACTGCTGTTTGGCCGACCGGACAGACCGCC 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCACGACAGAGTGTCCACTGTGCTCTCTGATCCAGGAGGT 1020
Db 2371 GCTGACTCCCATCCCTCTGGATCATGCAGGGTGTCCGCTGTGCTCTGATCCAGGAGGC 2430
QY 1031 ACCCATTCGCACCTCCCGATCAGGCTAAAGGCTTGCATGTTCTTCGATGGTAAAGTGC 1080
Db 2431 ACCCATTCGCCTCCCAATCGGGCTAAAGGCTTGCATGTTCTTCGATGGTAAAGTGC 2490
QY 1081 TGGGTTTCTCTAATAGAACTGAACACTGCTCAGTGGTTCACATGGTTCCTCTTCCATGAC 1140
Db 2491 TGGGTTTCTCTAATAGAGCTGAACACTGCTCAGTGGTTCACATGGTTCCTCTTCTGTGAC 2550
QY 1141 CCACGGCTTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTTGTA 1200
Db 2551 CCACAGCTTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTT-GAA 2609
QY 1201 TCTGTGAGGCCAAAGAACCCGAGTCAGAGAANGTGAGGCTTCCACCATTTGGGAAG 1257
Db 2610 TCCATAAGGCCAAAGAACCCGAGTCAGAGAACACGAGGCTTCCACCATTTGGGAG 2666

RESULT 19
AF52051382
LOCUS
DEFINITION
Homo sapiens individual 49 allele A, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION
AF520514
VERSION
AF520514.1 GI:33410996
KEYWORDS
2 of 2
SEGMENT
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2694)
AUTHORS
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED
14757826
REFERENCE
2 (bases 1 to 2694)
AUTHORS
Mallet,F., Bouton,O. and Oriol,G.
Direct Submission
JOURNAL
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
Location/Qualifiers
1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolates="individual 49 allele A"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="Asian
endogenous_virus: HERV-W"
43..57
/note="splice acceptor site"
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17579.1"
misc_feature
CDS
```

		/db_xref="GI:33410997"	
		/translation="MALPYHIFLFTVLLPSTLTAPPKRCMTSSSPYQFLWRMORP	
		GNIDAPYSRLSKGPTFTATHMPNCHYSATLCHNTHYWTGMNPSCFGILGV	
		TVCWTYFTQTMGGGVQDOAREKHKEVISQLTRVHGTSPPYKGLDLSKLHETLRT	
		HTRLVSLNLETHSNLTVCVKSNTYTTNSQIGIRWVTPQIVCLPSPGFVCGTSA	
		LVGPLVSNLETHSNLTVCVKSNTYTTNSQIGIRWVTPQIVCLPSPGFVCGTSA	
		YRLGNSSESMCFUSFLVPPMTIYTEODLYSVISKPRNKRVPILPVIAGVILGALG	
		TGIGITTSQFYKLSQELNGDMERVADSLVTLQDOLNSLAADVLRNRRALDLTAE	
		RGGTCFLGEBECYVNVQSGIVTEKVEIRDRQORAEELRNTGPMGLLSQWMPHLLP	
		FLGPLAAIILLPLPGPCLFNLLVNFVSRIEAVKLQMEPKMQSKTKIYRPLDRPASP	
		RSDVNDIKGTPPEEISAAQPLLRNSAGSS"	
		1705..1720	
		/note="splice acceptor site"	
		1959..2694	
		ORIGIN	
		Query Match 77.9%; Score 1035.2; DB 9; Length 2694;	
		Best Local Similarity 90.1%; Pred. No. 3e-304;	
		Matches 113; Conservative 0; Mismatches 111; Indels 13; Gaps 2;	
QY	1	TCAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGAACTGTTATTTT	60
Db	1423	TCAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAGAGGGGAACTGTTATTTT	1482
QY	61	AGGGAAGAAGTCTGTTAGTAGTATTAATCAATCTGGATCATTAATGAGAAAGTTAAAGA	120
Db	1483	AGGGAAGAAGTCTGTTATTAATTAATCAATCGGAATCGTCACTGAGAAGTTAAAGA	1542
QY	121	AATTTGAGATCGAATATTAATGAGAGAGAGACCTTCAAAACACTGCACCCCTGGGGCCT	180
Db	1543	AATTCGAGATCGAATACAAACGATAGAGAGAGAGGCTTCGAAACACTGAGCCCTGGGGCCT	1602
QY	181	CCTCAGCCAATGGATGGCCCTGGAGCTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT	240
Db	1603	CCTCAGCCAATGGATGGCCCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT	1662
QY	241	TTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTGTTAAGTTTGTCTTCCAGAAAT	300
Db	1663	GCTACTCTCTTTGGACCCCTGATCTTTAACTCTCTGTTAAGTTTGTCTTCCAGAAAT	1722
QY	301	TGAAGCTGTAAAGCTACAATAGTTCTTCAATGGAACCCCAAGATGCAAGTCCATGACTAA	360
Db	1723	CGAAGCTGTAAACTA-----CAATGAGGCCCAAGATGCAAGTCCAAGACTAA	1770
QY	361	AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTGATGCTCTGATGTTAATGACATTGA	420
Db	1771	GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA	1830
QY	421	AGTCACCCCTCCGAGGAATCTCAACTGCACAACCCCTACTACACTCCAAATTCAGTAGG	480
Db	1831	AGGCACCCCTCCTGAGGAATCTCAGCTGCACAACTTACTACGCCCAATTCAGCAGG	1890
QY	481	AAGCAGTTAGACAGTTGTTCAGCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540
Db	1891	AAGCAGTTAGAGGGTCTGTCGGCCAACTCCCAACAGCAGCTTAGGTTTCTGTTGAGA	1950
QY	541	GGGTGAGACTGAGAGACAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCNAGCCT	600
Db	1951	TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCTAAGCCT	2010
QY	601	ANCTGGGAAGGTGACCGCATCTTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660
Db	2011	AGCTGGGAAGGTGACCACTTCCACCTTTAAACAGGGGCTTGCAACTTAGCTCACACCTG	2070
QY	661	ACCAATCAGAGAGCTCACTAAAAATGCTAATCAGGCAAAAACAGGAGGTTAAGCAATAGCC	720
Db	2071	ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGCAGGAGTTAAGAAAATAGCC	2130
QY	721	AATCATCTATTGCTCAGAGGCAAGCGGGGAAGGACAGGATTTGGGATATAAACTCAGGCA	780
Db	2131	AATCATCTATTGCTCAGAGGCAAGCAGGAGGAGCAATGATCGGGATATAAAACCCAGTGC	2190
QY	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTTT	840

Db	2131	AATCACTATTGCTTGAGAGCAAGAGGAGGACAAATGATCGGATATAAACCCAAAGTC	2190
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTGCTCCCTCCCAATTGATGGGAGCTCTGTTTT	840
Db	2191	TTGAGCGGCAACAGGCAACCCCTTTGGGTGCTCCCTCCCTTTGATGGGAGCTCTGTTTT	2250
Qy	841	CACCTATTTCACCTCTATTAAATCATGCAACTGCACCTCTTCTGGTCCGTGTTTTTATGG	900
Db	2251	CATGCTATTTCACCTCTATTAAATCTTGCAACTGCACCTCTTCTGGTCCATGTTTCTTACGG	2310
Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTTTGCCACCGTCAAGACCCGCT	960
Db	2311	CTTGAGCTGAGCTTCGCTCGCCATCCACCACTGCTGTGTTTGCGCCACCGCAGACCCGCC	2370
Qy	961	GCTGACTTCCATCCCTTTGGATCAGCAGAGTGCTCACTGCTGCTCTGATCCAGCGAGGT	1020
Db	2371	GCTGACTTCCATCCCTCTGGATCATGACGGGTGTCCGCTGTGCTCTGATCCAGCGAGGC	2430
Qy	1021	ACCATTTGCCACTCCCGATCAGGCTAAAGGCTTGCCCAATGTTCTCGATGGCTAAGTGCC	1080
Db	2431	ACCATTTGCCGCTCCCAATCGGGCTAAAGGCTTGCCCAATGTTCTCGATGGCTAAGTGCC	2490
Qy	1081	TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTTCATGAC	1140
Db	2491	TGGGTTTCTTAATAGAGCTAATACACTACCGCATGGCCCAAGATTCCATTCCTTGTA	1200
Qy	1141	CCACGGCTTCTTAATAGAGCTAATACACTACCGCATGGCCCAAGATTCCATTCCTTGTA	1200
Db	2551	CCACAGCTTCTTAATAGAGCTAATACACTACCGCATGGCCCAAGATTCCATTCCTTGTA	2609
Qy	1201	TCTGTAGGCCCAAGAACCCGAGTGCAGAGANGTGGGCTTGGCCACATTTGGGAAG	1257
Db	2610	TCCATAGGCCCAAGAACCCGAGTGCAGAGANGTGGGCTTGGCCACATTTGGGAAG	2656
RESULT 21			
LOCUS	AF52051782	2594 bp	linear
DEFINITION	Homo sapiens individual 55 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.		
ACCESSION	AF520518		
VERSION	AF520518.1	GI:33411004	
KEYWORDS	2 of 2		
SEGMENT	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 2694)		
AUTHORS	Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.		
TITLE	The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
PUBMED	14757826		
REFERENCE	2 (bases 1 to 2694)		
AUTHORS	Mallet,F., Bouton,O. and Oriol,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France		
FEATURES	Location/Qualifiers		
source	1..2694		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/isolate="individual 55 allele A"		
	/db_xref="taxon:9606"		
	/chromosome="7"		
	/map="7q21-q22"		
	/sex="female"		
	/cell_type="PBMC's"		
	/note="endogenous_virus: HERV-W"		

misc_feature	43..57	/note="splice acceptor site"	
CDS	284..1900	/note="sncytin"	
		/codon_start=1	
		/product="envelope glycoprotein"	
		/protein_id="AAQ17581.1"	
		/db_xref="GI:33411005"	
		/translation="MALPYRHIFLPTVLPSFTLTAPPRCMTSSSPYQBFILWRMORP GNIDAPSYLSKGTPTTAHTHPRNCYSATLCMHANTHYTGKMINSPCGGLGV TWCVYFTQTGSDGGVQDQARKKHKEVLSQLTRVHGTSSTPYKGLDLKLKHEITLR HRLVLSNLEITLGHLEVAQNPNKVICLPNFRPVSIPVPEQWNNFSTEINTTSV LVGPLVSNLEITLGHLEVAQNPNKVICLPNFRPVSIPVPEQWNNFSTEINTTSV YCLGNSSSEMCTFLSPLVPMTIYTEQDLNYSIKPRNKRVPILPVRVIGAGVGLG YTGIGTSITPYKLSQELNGDMERVADSLVLIQDQNSLAIVLQNRRLALDLITAE RGCTCLFGBECCYVNGSIVTEKVEIRDRIQRRABEELRTPGMGLLSQWMPWILP FLGPLAAIILLFLFPCIFNLVNFVSRIEAVKLQMEPKVQSKTKIYRRLDRPASP RSDVNDIKGTPPEEISAAQLLRPNAGSS"	
misc_feature	1705..1720		
		/note="splice acceptor site"	
LTR	1359..2694		
ORIGIN			
	Query Match	77.9%; Score 1035.2; DB 9; Length 2694;	
	Best Local Similarity	90.1%; Pred. No. 3e-304;	
	Matches 1133; Conservative	0; Mismatches 111; Indels 13; Gaps 2;	
QY	1	TCMAAATCGAAGAGCTTTAGACTTCTCTAACCGCCAAAGAGAGGGGGAACCTGTTATTTTT	60
DB	1423	TCMAAATCGAAGAGCTTTAGACTTCTCTAACCGCTGAAGAGGGGGGAACCTGTTATTTTT	1482
QY	61	AGGGGAAGAATGCTGTAGTATGTTTAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA	120
DB	1483	AGGGGAAGAATGCTGTATTATGTTTAATCAATCCGANTCGTCACTGAGAAAGTTAAGA	1542
QY	121	AATTGAGATCGAATATAATATAGAGCAGAGAGACCTTCAAAACACTGACCCCTGGGCCT	180
DB	1543	AATTGAGATCGAATACAACTGAGAGCAGAGAGCTTCGAAACACTGGACCCCTGGGCCT	1602
QY	181	CCTCAGCCAAATGGAGCCCTGGACCTCCCTCTCTAGGACCTCTAGCAGCTATAATATT	240
DB	1603	CCTCAGCCAAATGGAGCCCTGGAGTCTCCCTCTCTAGGACCTCTAGCAGCTATAATATT	1662
QY	241	TTTACTCCTCTTTGGACCCCTGTATCTTCAACTTCTCTTTAAGTTTGTCTTCCAGAAAT	300
DB	1663	GCTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCTCTTGAACCTTGTCTTCCAGAAAT	1722
QY	301	TGAAGCTGTAAAGCTACAAATAGTTCCTCAATATGAACCCAGATGCGAGTCCATGACTAA	360
DB	1723	CGAAGCTGTAAAGCTA-----CAATGGAGCCCAAGATGCGAGTCCAGACTAA	1770
QY	361	AATCTACCTGGACCCCTGGACCGGCCTGTAGACTATGCTCTGATGTTAATGACTGA	420
DB	1771	GATCTACCGCAGACCCCTTGAACCGGCCTGTAGCCACGATCTGATGTTAATGACTCA	1830
QY	421	AGTCAACCCCTCCGAGGAATCTCAATGTGCAACACCCCTACTACACTCCAATTCAGTAGG	480
DB	1831	AGGCACCCCTCTCTGAGGAATCTCAGCTGSCACAACCTCTACTACGCCCAATTCAGCAGG	1890
QY	481	AAGCAGTTAGACAGTTGTGCAGCAACCTTCCCAACAGTACTTGGGTTTTCTCTTCGAGA	540
DB	1891	AAGCAGTTAGACAGTTGTGCAGCAACCTTCCCAACAGTACTTGGGTTTTCTCTTCGAGA	1950
QY	541	GGGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCNAAAGCT	600
DB	1951	TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTTAAGCCT	2010
QY	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGCACCTTAGCTCACACCG	660
DB	2011	AGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGCACCTTAGCTCACACCTG	2070
QY	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCACAAAACAGGAGGTAAAGCAATAGCC	720

QY	661	ACCAATCAGAGAGCTCACATAAATGCTTAATCAGGCCAAAACAAGAGGTAAAGCAATAGCC	720
Db	2071	ACCAATCAGAGAGCTCACATAAATGCTTAATTAGGC AAAACAAGAGGTAAAGCAATAGCC	2130
QY	721	AATCATCTATTGGCTTGAGAGCACAGCGGGAAGGACAAGGATTCGGATATATAAATCTCAGGCA	780
Db	2131	AATCATCTATTGGCTTGAGAGCACAGCGGGAAGGACAATGATCGGATATATAAATCTCAGGCA	2190
QY	781	TTCAGCGCAGCAACAGCAACCCCCTTTGGGTCCCTCCCAATGATTCGGAGTCTGTGTTTT	840
Db	2191	TTCAGCGCAGCAACAGCAACCCCCTTTGGGTCCCTCCCAATGATTCGGAGTCTGTGTTTT	2250
QY	841	CACCTATTTCACCTCTATTAAATCATGCAACTGCACCTCTTCTCGTCGGTCTTTTTTATGG	900
Db	2251	CATGCTATTTCACCTCTATTAAATCATGCAACTGCACCTCTTCTCGTCGGTCTTTTTTATGG	2310
QY	901	CTCAAGCTGAGCTTTTGTTCGCGCATCACCACTGCTGTTTGGCCACCCTGCACAGACCCGCT	960
Db	2311	CTTGAGCTGAGCTTTGCTGCGCATCCACACTGCTGTTTGGCCACCCTGCACAGACCCGCT	2370
QY	961	GCTGACTTCCATCCCCCTTTGGATCCAGAGAGTGTCCA CTGTGCTCTGTATCCAGCGAGGT	1020
Db	2371	GCTGACTTCCATCCCCCTTTGGATCCAGAGAGTGTCCA CTGTGCTCTGTATCCAGCGAGGT	2430
QY	1021	ACCAATGACACTCCCGATCAGGCTAAAGGCTTGGCCATCTCTCTGCATCGCTAAGTGCC	1080
Db	2431	ACCAATGACACTCCCGATCAGGCTAAAGGCTTGGCCATCTCTCTGCATCGCTAAGTGCC	2490
QY	1081	TGGTTTGTCTTAATAGA ACTGAACTGCTGCTACTGGGTTCCATGGTTCTCTCCATGAC	1140
Db	2491	TGGTTTGTCTTAATAGA ACTGAACTGCTGCTACTGGGTTCCATGGTTCTCTCCATGAC	2550
QY	1141	CCACGGCTTCTAATAGAGCTATAACACTACCGCATCGCGATCGCCCAAGTTCATTCCTGGTA	1200
Db	2551	CCACGGCTTCTAATAGAGCTATAACACTACCGCATCGCGATCGCCCAAGTTCATTCCTGGTA	2610
QY	1201	TCTGTGAGGCAAGAACCCAGGTCAGAGANGTGAGGCTTGCACCATTTGGGAAG	1257
Db	2610	TCCATAAGGCAAGAACCCAGGTCAGAGANGTGAGGCTTGCACCATTTGGGAAG	2666
RESULT 23			
AF520521S2			
LOCUS			
DEFINITION Homo sapiens individual 6 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.			
ACCESSION AF520522			
VERSION AF520522.1 GI:33411012			
KEYWORDS			
SEGMENT			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			


```
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/notes="Caucasian
endogenous_virus: HERV-W"
misc_feature
43..57
/notes="splice acceptor site"
284..1900
/notes="synctin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17583.1"
/db_xref="GI:33411013"
translation="MALPYHIFLFTVLPSFTLPAFPCRCMTSSSPYQEFLLMRQRP
GNIDAPYSRLSKGTPTPTAHTHPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGV
TIVTYSFTQMGSGGGVQDQAREKHKEVISLQTRVHGTSSPYKGLDLSKLHETLRT
HTRLVSLPNTTLTGLHVSACNPNCWICLPLNRPVYSIIPVROMNFSFEINTTSS
LVGLVSNLEITHRSNLTKVFSNTTNTNSQCIRWVTPPTQIVCLPSGIFPVCGTSA
YRCUNGSESCFLSFLVPPMTIITEQDLINYIVISKRNKRVPILPFVIGVGLGALG
TGIGITSTQTFYKLSQELNGDMERVAESLVTLOQLNSIAAVLQNRALDILLTAE
RGGCLFLGECCEYVNGSIGIVTEKVEIRDIQIRABELNRTGFWLLSQMFWIIP
FLGPLAAIILLLLPGPCIFNLVNFVSRISIEAVKLQMEPKMQSKTKIYRPLDRPAS
RSDVNDIKGTPPEISAAQPLLRPNASGSS"
misc_feature
1705..1720
/notes="splice acceptor site"
1959..2694
ORIGIN
```

```
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304; Indels 13; Gaps 2;
Matches 1133; Conservative 0; Mismatches 111;

Qy 1 TC AAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGAGGGGAACTGTGTTATTTT 60
Db 1423 TC AAAATCGAAGAGCTTTAGACTTCTAAACCGCTGAAAGAGGGGAACTGTGTTATTTT 1482

Qy 61 AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCATTAAGAGAAAGTTAAGA 120
Db 1483 AGGGGAAGAATGCTGTTATGTTAATCAATCTCGGAATCGTCACTGAGAAAGTTAAGA 1542

Qy 121 AATTGAGATCGAATATAATGATAGAGCAGAGGACCTTCAAAACACCTGCACCCCTGGGGCT 180
Db 1543 AATTCGAGATCGAATACACGATGAGCAGAGGAGCTTCGAAACACTGACCCCTGGGGCT 1602

Qy 181 CCTACGCAATGGATGCGCTCGACTCTCCCTTTCTAGGACTCTTAGCAGCTATAATATT 240
Db 1603 CCTACGCAATGGATGCGCTCGACTCTCCCTTTCTAGGACTCTTAGCAGCTATAATATT 1662

Qy 241 TTTACTCCTCTTTGGACCGCTGATCTTCAACTTCCTCTGTTAAGTTTGTCTCTCCAGAA 300
Db 1663 GCTACTCCTCTTTGGACCGCTGATCTTAAACCTCCTCTGTTAAGTTTGTCTCTCCAGAA 1722

Qy 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCGAGTCCATGACTAA 360
Db 1723 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCGAGTCCAAGACTAA 1770

Qy 361 AATTCACGCTGGAACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 1771 GATCTACGCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830

Qy 421 AGTCACCCCTCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
Db 1831 AGGCACCCCTCTCTGAGGAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 1890

Qy 481 AAGCAGTTAGACAGTTGTGACGCAACCTCCCCAACAGTACTTGGGTTTTCTGTGTAGA 540
Db 1891 AAGCAGTTAGACGCTCGTGGCCCAACCTCCCCAACAGCAGCTTAGGTTTTCTGTGTAGA 1950

Qy 541 GGGTGGACTGAGAGACAGACTAGCTGGAATTCCTAGGCTGACTAAGAAATCCCNAGGCT 600
Db 1951 TGGGGGACTGAGAGACAGGACTAGCTGGAATTCCTAGGCTGACTAAGAAATCCCNAGGCT 2010
```

```
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTTGCAACTTAGCTCACACCG 660
Db 2011 AGCTGGGAAGGTGACCGCATCCACCTTTTAAACACGGGGCTTTGCAACTTAGCTCACACCTG 2070

Qy 661 ACCAATCAGAGAGCTCAATAAATGCTAAATCAGGCAAAAAACAGGAGGTAAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGCTCAATAAATGCTAAATAGGCAAAAAACAGGAGGTAAAGCAATAGCC 2130

Qy 721 AATCATCTATTGCTCTGAGAGCAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA 780
Db 2131 AATCATCTATTGCTCTGAGAGCAGCAGGAGGGAACAATGATCGGGATATAAATCAGGCA 2190

Qy 781 TTCAGGCGCAACACAGCAACCCCTTTGGGTCCCTCCCATTTGATTTGATGGAGCTCTGTTTT 840
Db 2191 TTCAGGCGGCAACAGGCAACCCCTTTGGGTCCCTCCCTTTGATTTGATGGAGCTCTGTTTT 2250

Qy 841 CACTCTATTTTCACTCTATTAAATCATCAACTGCATCTCTTCTGTCGGTGTGTTTTATGG 900
Db 2251 CATGCTATTTTCACTCTATTAAATCTTGCAACTGCATCTCTTCTGTCATGTTTCTTACGG 2310

Qy 901 CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTCTGCTCGCCATCCACCACTGCTGTTTGGCCACCGTCAGACCCGCC 2370

Qy 961 GCTGACTTTCATCCCTTTGGATCCAGCAGAGTGTCCTGCTGCTCTGATCCAGCGAGGT 1020
Db 2371 GCTGACTTTCATCCCTCTGATCATGCAAGGTGTCGCTGCTCTGATCCAGCGAGGC 2430

Qy 1021 ACCCATTTGCCACTCCCGATCAGGCTAAAGGCTTGCCATTTGTTCTCTGATGCTGAGTCC 1080
Db 2431 ACCCATTTGCCGCTCCCAATTTGGGCTTAAAGGCTTGCCATTTGTTCTGATGCTGAGTCC 2490

Qy 1081 TGGGTTTTGCTCTATAGAACTGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 2491 TGGGTTTTGCTCTATTAAGCTGAACACTAGTCACTGGGTTCCATGTTCTCTTCTGTCGAC 2550

Qy 1141 CCAGCGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTTCCTTGTA 1200
Db 2551 CCACAGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTTCCTT-GAA 2609

Qy 1201 TCTGTAGGCGCAAGAACCCCGATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 1257
Db 2610 TCCATAAGGCGCAAGAACCCCGATGAGGAGTGTGAGGAGTGTGAGGAGTGTGAGGAG 2666

RESULT 24
AF52052582 2694 bp DNA linear PRI 11-FEB-2004
LOCUS Homo sapiens individual 65 allele A, envelope glycoprotein gene,
DEFINITION complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520526
VERSION AF520526.1 GI:33411020
KEYWORDS
2 of 2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2694)
AUTHORS Mallet,F., Lucotte,G., Duret,L. and Mandrand,B.
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
JOURNAL 14757826
PUBMED 2 (bases 1 to 2694)
REFERENCE Mallet,F., Bouton,O. and Oriol,G.
AUTHORS Direct Submission
TITLE Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
JOURNAL d'Italie, Lyon 69364 cedex 07, France
FEATURES Location/Qualifiers
```


1771	GATCTACCGCAGACCCCTGGACCGGCTCTAGTACGCCACAGATCTGTATGTAAATGACATCAA	1831
421	AGTCAACCCCTCCCGAGGAATCTCAACTGCAGAAACCCCTACTACACTTCCAATTCAGTAGG	480
1831	AGGCACCCCTCTCTGAGGAATCTCAGCTGCACAACCTCTACTAGCCCCCAATTCAGCAGG	1890
481	AAGCAGTTAGACGAGTTGTTCAGCAACCTCCCCCAACAGTACTTTGGGTTTTCTCTGTTGAGA	540
1891	AAGCAGTTAGAGGGTCTGTTCGCGCAACCTCCCCCAACAGCAGTCTTGGTTTTCTCTGTTGAGA	1950
541	GGTGGACTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	600
1951	TGGGGGACTAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTTAAGCCT	2010
601	ANCTGGGAAGGTGACCGGATCCATCTTTAAACATGGGGCTTGCACCTTAGCTACACCCG	660
2011	AGCTGGGAAGGTGACCGGATCCATCTTTAAACATGGGGCTTGCACCTTAGCTACACCCG	2070
661	ACCAATACAGAGACTCAGCTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC	720
2071	ACCAATACAGAGACTCAGCTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC	2130
721	AATCATCTAATTCCTGAGAGCAGCGGAAGGACAAGGATTTGGGATATAAACTCAGGCA	780
2131	AATCATCTAATTCCTGAGAGCAGCGGAAGGACAAGGATTTGGGATATAAACTCAGGCA	2190
781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTT	840
2191	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTT	2250
841	CAGTCTAATTTCACTCTAATTAATCATGCAACTGCAGCTCTTCTGCTCGTGTTTTTATGG	900
2251	CAGTCTAATTTCACTCTAATTAATCATGCAACTGCAGCTCTTCTGCTCGTGTTTTTATGG	2310
901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGCTCAGAGCCCGCT	960
2311	CTTGAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGCTCAGAGCCCGCT	2370
961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCAAGTGTCTGCTCTGATCCAGGAGGT	1020
2371	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCAAGTGTCTGCTCTGATCCAGGAGGT	2430
1021	ACCATTTGCCATCTCCGATCAGGCTAAGGCTTGGCAATCTTCTGCTGATGGCTAAGTGCC	1080
2431	ACCATTTGGCGCTCCCAATCGGCTAAGGCTTGGCAATCTTCTGCTGATGGCTAAGTGCC	2490
1081	TGGGTTTGTCTAATAGAACTGAAACCTGCTCACTGGGTTTCCATGGTTCTCTCCATGAC	1140
2491	TGGGTTTGTCTAATAGAACTGAAACCTGCTCACTGGGTTTCCATGGTTCTCTCCATGAC	2550
1141	CCACGGCTTCTAATAGAGCTATAACACTACCGCATGGCCCAAGATTCATCTCTTGGTA	1200
2551	CCACGGCTTCTAATAGAGCTATAACACTACCGCATGGCCCAAGATTCATCTCTTGGTA	2610
1201	TCGTGAGGCCAAGAACCCAGGCTCAGAGAAAGTGGGCTTGCACCATTTGGGAAG	1257
2610	TCCATAAGGCCAAGAACCCAGGCTCAGAGAAAGTGGGCTTGCACCATTTGGGAAG	2666
RESULT 27		
AF520537S2		
LOCUS	2694 bp DNA linear	PRI 11-FEB-2004
DEFINITION	Homo sapiens individual 81 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.	
ACCESSION	AF520538	
VERSION	AF520538.1	GI:33411044
KEYWORDS		
SEGMENT	1	
SOURCE	2 of 2	
ORGANISM	Homo sapiens (human)	
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2694)	

```
AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G.,
Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.
TITLE The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED 14757826
REFERENCE 2 (bases 1 to 2694)
AUTHORS Mallet,F., Bouton,O. and Oriol,G.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
    source
        1..2694
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /isolate="individual 81 allele A"
            /db_xref="taxon:9606"
            /chromosome="7"
            /map="7q21-q22"
            /sex="female"
            /cell_type="PBMC's"
            /note="endogenous_virus: HERV-W"
            /note="splice acceptor site"
            43..57
                misc_feature
            284..1900
                CDS
                /note="syncytin"
                /codon_start=1
                /product="envelope glycoprotein"
                /protein_id="AAQ17591.1"
                /db_xref="GI:33411045"
                /translation="MALPYHFLFTVLPLSFTLTAPPCRCMTSSSPVOEPLWRMQRP
                GNIDAPSVRSLSKGTPTTAHTMPRNCYHSATLCMHANTHYWTGKMINPSCPGSLGV
                TVCTYPTQTSMDGGVQDQAREKHVKEVISLQTRVHGTSPPYKGLDLSKLHETLR
                HRLVSLFTTLGHEVSAQNPNCWICLPLNRPYVISIPVPMQNFSTEINTSV
                LVGLVSNLEITHTSNLTCVRFESNTTYTNSQCIWRWTPPTQIVCLPSGIFVFCGTA
                YRCLNGSESSECFSLFVPPMTIYEQDLYSVISKPRKRPVILPFIAGVILGALG
                TGIGITTFQFYKLSQELNGDMERVADSLVTLQDLNSLAAVVLQNRALDILLTAE
                RGGTCLFLGECCCVYVNGSIGIVTEKVEIRDIORRAEELRNTGPGWLLSOWMPWILP
                FLGPLAAILLLLPGPCIFNLLVPVSRIEAVKLQMPKQSKTKIYRRPLDRPAS
                RSDVNDIKGTPEEISAAQPLLRPNAGSS"
                1705..1720
                    misc_feature
                /note="splice acceptor site"
                1959..2694
                    LTR
                    ORIGIN
Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
Qy 1 TCAAATCGAAGAGCTTTAGACTTGCTAACCGCAAAAGAGGGGAACCTGTTATTATTTT 60
Db 1423 TCAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGGGGAACCTGTTATTATTTT 1482
Qy 61 AGGGAAGAATGCTTTAGTATGTTAATCAATCTGGAATCAATTACTGAGAAGTTAAAGA 120
Db 1483 AGGGAAGAATGCTTTAGTATGTTAATCAATCGGAATCGTCACTGAGAAGTTAAAGA 1542
Qy 121 AATTGAGATCGAATATAATATGATAGAGAGAGAGACCTTCAAAACACTGCACCTGGGGCT 180
Db 1543 AATTGAGATCGAATATAACGCTAGAGAGAGAGAGCTTCGAAACACTGCACCTGGGGCT 1602
Qy 181 CCTAGCAATGAGTGCCTGACTCTCCCTTCTTAGGACTCTTAGCAGCTATAATATT 240
Db 1603 CCTAGCAATGAGTGCCTGAGTTCTCCCTTCTTAGGACTCTTAGCAGCTATAATATT 1662
Qy 241 TTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGAAGTTTGCTCTTCCAGAAAT 300
Db 1663 GCTACTCTCTTTGGACCTGTATCTTTAACTCTCTTGAAGTTTGCTCTTCCAGAAAT 1722
Qy 301 TGAAGCTGAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCGAGTCATGACTAA 360
Db 1723 CGAAGCTGTAATAACTA-----CAATGGAGCCCAAGATGCGAGTCATGACTAA 1770
```

```
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 1771 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCAA 1830
Qy 421 AGTCACCCCTCCCGAGGAAATCTCAATGTCACAAACCCCTACTACACTCTCAATTCAGTAGG 480
Db 1831 AGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG 1890
Qy 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTAGTCTGGGTTTCTCTGTTGAGA 540
Db 1891 AAGCAGTTAGAGCGGTGCTGCGGCAACCTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGA 1950
Qy 541 GGGTGGACTGAGACAGGAGCTAGCTGGATTTCCTAGGCTGACTAAGAATCCCAAGCCT 600
Db 1951 TGGGGGACTGAGACAGGAGCTAGCTGGATTTCCTAGGCTGACTAAGAATCCCAAGCCT 2010
Qy 601 ANCTGGGAAGGTGACCGATCTCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 2011 AGCTGGGAAGGTGACCATCTCCCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2070
Qy 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGTTAAAGCAATAGCC 720
Db 2071 ACCAATCAGAGAGCTCACTAAATGCTAATTAGCAAAAGACAGGAGTTAAAGCAATAGCC 2130
Qy 721 AATCATCTATTGCTGAGAGCACAGCGGAGGACAAAGGATTGGGATATAACTCAGGCA 780
Db 2131 AATCATCTATTGCTGAGAGCACAGCAGGAGGACAAATGATCGGATATAAACCAAGTC 2190
Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCTCCATTTGATATGGAGCTCTGTTTT 840
Db 2191 TTCAAGCAGGCAACAGCAACCCCTTTGGGTCCCTCTCCCTTTGATATGGAGCTCTGTTTT 2250
Qy 841 CACTCTATTTCACTCTATTAATCATCACTGACACTCTCTGCTCGTGTGTTTTTATGG 900
Db 2251 CATGCTATTTCACTCTATTAATCTTGCAACTGACACTCTCTGCTCCATGTTTCTTACGG 2310
Qy 901 CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAAGACCCGCT 960
Db 2311 CTTGAGCTGAGCTTTTCGCTCGCCATCCACCACCTGCTGTTTGGCCGCCACCGCAGACCCGCC 2370
Qy 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGGAGGT 1020
Db 2371 GCTGACTCCCATCCCTCTGATCATGACGGGTGTCCGCTGTGCTCTGATCCAGCGAGGC 2430
Qy 1021 ACCCATTGGCCACTCCCGATCAGGCTAAAGGCTTCGCCATTGTTCTGTCATGCTAAGTGCC 1080
Db 2431 ACCATTGGCGCTCCCAATCGGGCTAAAGGCTTGCCATTGTTCTGTCATGCTAAGTGCC 2490
Qy 1081 TGGGTTTGTCTTAATAGAACTGAACACTGTGCTCACTGGGTTCCATGGTTCTCTTCCATGAC 1140
Db 2491 TGGGTTTCATCTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGAC 2550
Qy 1141 CCAGCGGCTTCTAATAGAGCTATAACACTCACCGCATGCCCCCAAGATTCCTTCTTGTGTA 1200
Db 2551 CCACAGCTTCTAATAGAGCTATAACACTCACCGCATGCCCCCAAGGTTCCATTCCTT-GAA 2609
Qy 1201 TCTGTGAGGCCCAAGAACCCCGAGTTCAGAGAAGTGTAGGCTTGCCACCAATTTGGGAAG 1257
Db 2610 TCCATAAGGCCCAAGAACCCCGAGTTCAGAGAAGTGTAGGCTTGCCACCAATTTGGGAG 2666
RESULT 28
AF520549S2
LOCUS AF520549S2 2694 bp DNA linear PRI 11-FEB-2004
DEFINITION Homo sapiens individual 95 allele A, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
ACCESSION AF520550
VERSION AF520550.1 GI:33411068
KEYWORDS
SEGMENT 2 of 2
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2694)
Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.

The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology

Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)

JOURNAL

PUBMED

REFERENCE

Mallet,F., Bouton,O. and Oriol,G.

TITLE

Direct Submission

JOURNAL

Submitted (07-JUN-2002) Retrovirology Department, UMR 2142

CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée

d'Italie, Lyon 69364 cedex 07, France

FEATURES

source

1..2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 95 allele A"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q21-q22"
/sex="female"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
misc_feature
43..57
/notes="splice acceptor site"
284..1900
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA017597.1"
/db_xref="GI:33411069"
/translation="MALPYHFLFTVLLPSFTLTAPPCCRMSTSSPYQFELWRMQR
TNDAPYSLSKGTFTFTAHPRNCVHSATLCMHANTHYMTGMINPSPCGGLGV
TWCTVYTGSGDGVODQAREKHVEVLSQTRVHGTSSPYKGLDLSKLHETRT
HRLVSLFWTLTGLHEVSAQNPNCWICLPNFRPVYSIPVEOWNFESTINTSV
LVGLVSLNEIHTISNLTCKFSNTITYSQCIRWVPTQIVCLPSGIFPVCGTSA
YKLGSSSESMFLSLVPPMTIYTEQDLYSVISKPRNRPVPLFPFVIGVGLAG
TGIGTITTFQYKLSQBLNGDMERVADSLVLOQLNSLAALVQLNRALDLTAE
RGCTCLFTEGCVYVQSGIVTEKVEIKRDRIQRABELNTGPMGLLSQWMPWILP
FLGPLAAIILLLLFGPCIENLLVNFVSSRIEAVKLQWEPKMQSKTIYRRLDRPASP
RSDVNDIKGTPEEISAAQPLRPNSAGSS"
misc_feature
1705..1720
/notes="splice acceptor site"
1959..2694

CDS

Query Match 77.9%; Score 1035.2; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 3e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;
ORIGIN
1 TCAAAATCGAAGAGCTTTAGACTTGCTAACCGCAAAAGAGGGGAAACCTGTTATTTT 60
1423 TCAAAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAGAGGGGAAACCTGTTATTTT 1482
61 AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATCTAGAGAAAGTTAAAGA 120
1483 AGGGGAAGAAATGCTGTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542
121 AATTGAGATCGAATATAAATGAGCAGGAGGACCTTCAAAACACTGCACCCCTGGGCGCT 180
1543 AATTGAGATCGAATACAACGTAGCAGAGGAGCTTCGAAACACTGGACCCCTGGGCGCT 1602
181 CCTCAGCAATGATGCTCCCTGGACTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT 240
1603 CCTCAGCAATGATGCTCCCTGGACTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT 1662
241 TTTTACTCTCTTTGGACCCCTGTTATCTTCAACTCTCTTTAGTTGTTCTCTTCCAGAAAT 300
1663 GCTACTCTCTTTGGACCCCTGTTATCTTAACTCTCTTGTAACTTTGTTCTCTTCCAGAAAT 1722

301 TGAAGCTGTAAGCTTACAAATAGTCTTTCAATAGGAACCCAGATGCACTCAATCACTAA 360
1723 CGAAGCTGTAAGCTTACAAATAGTCTTTCAATAGGAACCCAGATGCACTCAATCACTAA 1770
361 AATCTACCGTGGAGCCCTTGGACCGGCTGTAGACTATGCTCTGATGTATGATGCAATTGA 420
1771 GATCTACCGCAGACCCCTTGGACCGGCTGTAGCCACGATCTGATGTATGATGCAATCAA 1830
421 AGTCACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTACTACTCAATTCAGTAGG 480
1831 AGGACCCCTCTCTGAGGAAATCTAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 1890
481 AAGCAGTTAGACAGCTTGTCTAGCCAACTCTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
1891 AAGCAGTTAGAGCGTCTGCGGCCAACTCTCCCAACAGCAGCTTAGGTTTTCTGTTGAGA 1950
541 GGTGGAGCTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGGATCCCNAGCCT 600
1951 TGGGGGACTGAGAGACAGGACTAGCTGATTTCTTAGCTGACTAGGATCCCNAGCCT 2010
601 ANCTGGGAAGGTGACCGCATCTCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660
2011 AGCTGGGAAGGTGACCATCTCCCTTTAAACAGCGGGCTTGCAACTTAGCTCACACCTG 2070
661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
2071 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
721 AATCATCTATTGCTTGGAGGACAGCGGAGGACAGGATTTGGATATAAATCACTCAGGCA 780
2131 AATCATCTATTGCTTGGAGGACAGCGGAGGACAGGATTTGGATATAAATCACTCAGGCA 2190
781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGTTTT 840
2191 TTGAGCGCGCAACCGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGTTTT 2250
841 CACTCTATTTCACCTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTGTTTTATGG 900
2251 CATGCTATTTCACCTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTACGG 2310
901 CTCAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTGTCACCGTCACAGCCCGCT 960
2311 CTTGAGCTGAGCTTTGCTGCGCATCCCACTGCTGTTGTCGCGCACCCGACAGCCGCC 2370
961 GCTGACTTCCATCCCTTTGGATCCAGAGGTGTCACCTGCTGCTCTGATCCAGCGAGGT 1020
2371 GCTGACTTCCATCCCTTCTGGATCATGAGGGTGTCCGCTGCTGCTGATCCAGCGAGG 2430
1021 ACCCATTTGCCACTCCCGATCAGGCTAAAGCTTGGCATTTGCTCTGATGGCTAAGTGCC 1080
2431 ACCCATTTGCCCTCCCAATCGGGCTAAAGGCTTGGCATTTGCTCTGATGGCTAAGTGCC 2490
1081 TGGTTTTGCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
2491 TGGGTTTCATCTTAATAGAGCTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550
1141 CCACGGCTTCTAATAGAGCTATAAACAACCTCAGCGATGGCCCAAGATTTCATTTCTGGTA 1200
2551 CCACAGCTTCTAATAGAGCTATAAACAACCTCAGCGATGGCCCAAGATTTCATTTCTGTA 2609
1201 TCTGTGAGGCCAAGAACCCAGGTCAGAGAANGTAGGCTTGCACCATTTGGGAAG 1257
2610 TCCATAAGGCCAAGAACCCAGGTCAGAGAACACGAGGCTTGCACCATTTGGGAG 2666

RESULT 29

AF520553S2

LOCUS

DEFINITION

AF520554

AF520554.1

GI:33411076

KEYWORDS

2694 bp DNA linear PRI 11-FEB-2004
Homo sapiens individual 96 allele A, envelope glycoprotein gene,
complete cds, and 3' long terminal repeat, complete sequence.
AF520554
AF520554.1

[illegible]

ACCESSION	AF520562	Db	1603	CCTCAGCCAAATGGATGCGCTGCTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT	1662
VERSION	AF520562.1				
KEYWORDS	2 of 2	Qy	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTCTGTAGTTTGTCTCTTCCAGAAAT	300
SEGMENT		Db	1663	GCTACTCTCTTTGGACCTGTATCTTAAACCTCTCTGTAACTTGTCTCTTCCAGAAAT	1722
SOURCE	Homo sapiens (human)	Qy	301	TGAAGCTGTAAAGCTTACAAATAGTCTTCAAATGGAAACCCAGATGCGAGTCCATGACTAA	360
ORGANISM	Homo sapiens	Db	1733	CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCGAGTCCCAAGACTAA	1770
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Qy	361	AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGTCTGTATGTAAATGACATTGA	420
AUTHORS	Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.	Db	1771	GATCTACCGCAGACCCCTGGACCGGCTGTAGCCAGCATCTGATGTAAATGACATCAA	1830
TITLE	The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology	Qy	421	AGTCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATCAGTAGG	480
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)	Db	1831	AGGCACCCCTCTCTGAGGAAATCTAGCTGCACAACTCTACTACGCCCAATTCAGCAGG	1890
PUBMED	14757826	Qy	481	AAGCAGTTAGAGCAGTGTGTACGCCAACTCCCAACAGTACTTGGGTTTTCTCTTTGAGA	540
REFERENCE	2 (bases 1 to 2694)	Db	1891	AAGCAGTTAGAGCGGTGCTGCGCCAACTCCCAACAGCAGTACTTGGGTTTTCTCTTTGAGA	1950
AUTHORS	Mallet,F., Bouton,O. and Oriol,G.	Qy	541	GGGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT	600
TITLE	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-biomerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France	Db	1951	TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT	2010
JOURNAL	Location/Qualifiers	Qy	601	ANCTGGGAGGTGACCGGCTTCCATCTTTAAACATCGGGCTTGCACCTTAGCTAGCTCACACCGG	660
FEATURES	1. .2694	Db	2011	AGCTGGGAGGTGACACATCCACTTTAAACACGGGGCTTGCACCTTAGCTAGCTCACACCTG	2070
source	/organism="Homo sapiens" /mol_type="genomic DNA" /isolates="individual 71 allele A" /db_xref="taxon:9606" /chromosome="7" /map="7q21-q22" /sex="male" /cell_type="PBMC's" /notes="African endogenous_virus: HERV-W"	Qy	661	ACCAATCAGAGAGCTCATTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	720
misc_feature	43. .57	Db	2071	ACCAATCAGAGAGCTCATTAAATGCTAATTAGGCAAAACAGGAGGTAAAGAAATAGCC	2130
CDS	284. .1900	Qy	721	AATCATCTATCCCTGAGAGCAGCGGGAAGGACAGGATTTGGGATATAAACTCAGGCA	780
	/codon_start=1 /product="envelope glycoprotein" /protein_id="AAQ17603.1" /db_xref="GI:33411093" /translation="MALPYHFLFTLLPSFTLTAPPCKMTSSSPYOEFLWRMORP GNIDARSLSKGTPTFTAHMPNCVHSATLQNHANTHYWTGWINPSPGVLGV HTVLVLFNTTLGLHEVDAQREKHVEVISQLTRVHGTSSPYKGLDLSKLHETLT LVGLVSNLEITHTSNLTCKVFSNTTNTNSQCIRWTPPTQIVCLPSGIFVCGTSA YRLGSSSSEMFSLPVPMTIYTDQDLYNYVISKPRNRPVILPFIAGVIGLALG TGIGGTTSTQFYKLSOELNGDMERVADSLVTDQLNSLAALAVLQNRALDLTAE RGTCIFLGECCYYNQSIGIVTEKVEIRDRIOREARELRNTPKGLLSQMPWILP FLGPLAAILLLFLGFCIFNLLNVFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP RSDVNDIKGTPPEEISAAQPLLRPNAGSS"	Db	2131	AATCATCTATTCCTGAGAGCAGCAGGAGGACAAATGATCGGGATATAAAACCCAGTC	2190
	/note="syncytin"	Qy	781	TTCAAGCCAGCAACACACACCCCTTTGGGTCCCTCCCATTTGTATGGGAGCTCTGTTTT	840
	43. .57	Db	2191	TTGAGCCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTT	2250
	284. .1900	Qy	841	CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGCTGTTTTATGG	900
	/product="envelope glycoprotein" /protein_id="AAQ17603.1" /db_xref="GI:33411093" /translation="MALPYHFLFTLLPSFTLTAPPCKMTSSSPYOEFLWRMORP GNIDARSLSKGTPTFTAHMPNCVHSATLQNHANTHYWTGWINPSPGVLGV HTVLVLFNTTLGLHEVDAQREKHVEVISQLTRVHGTSSPYKGLDLSKLHETLT LVGLVSNLEITHTSNLTCKVFSNTTNTNSQCIRWTPPTQIVCLPSGIFVCGTSA YRLGSSSSEMFSLPVPMTIYTDQDLYNYVISKPRNRPVILPFIAGVIGLALG TGIGGTTSTQFYKLSOELNGDMERVADSLVTDQLNSLAALAVLQNRALDLTAE RGTCIFLGECCYYNQSIGIVTEKVEIRDRIOREARELRNTPKGLLSQMPWILP FLGPLAAILLLFLGFCIFNLLNVFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP RSDVNDIKGTPPEEISAAQPLLRPNAGSS"	Db	2251	CATGCTATTTCATCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACAG	2310
	43. .57	Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCAACAATGCTGTTTGGCCCGTCAACAGCCGCT	960
	284. .1900	Db	2311	CTTGAGCTGAGCTTTTCGCTCGCCATCCACACTGCTGTTTGGCCCGCACCGCAGCCGCC	2370
	/note="syncytin"	Qy	961	GCTGACTTCATCCCTTTGGATCCAGCAGTGTCCACTGTGCTCTCTGATCCAGCCGAGGT	1020
	43. .57	Db	2371	GCTGACTTCATCCCTCTGGATCAGCAGGCTGCTCGCTGTGCTCTCTGATCCAGCCGAGGC	2430
	284. .1900	Qy	1021	ACCCATTGCACTCCCGATCAGGCTAAAGCTTGCATTTGCTTCTGATGGCTTAAGTGCC	1080
	/product="envelope glycoprotein" /protein_id="AAQ17603.1" /db_xref="GI:33411093" /translation="MALPYHFLFTLLPSFTLTAPPCKMTSSSPYOEFLWRMORP GNIDARSLSKGTPTFTAHMPNCVHSATLQNHANTHYWTGWINPSPGVLGV HTVLVLFNTTLGLHEVDAQREKHVEVISQLTRVHGTSSPYKGLDLSKLHETLT LVGLVSNLEITHTSNLTCKVFSNTTNTNSQCIRWTPPTQIVCLPSGIFVCGTSA YRLGSSSSEMFSLPVPMTIYTDQDLYNYVISKPRNRPVILPFIAGVIGLALG TGIGGTTSTQFYKLSOELNGDMERVADSLVTDQLNSLAALAVLQNRALDLTAE RGTCIFLGECCYYNQSIGIVTEKVEIRDRIOREARELRNTPKGLLSQMPWILP FLGPLAAILLLFLGFCIFNLLNVFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP RSDVNDIKGTPPEEISAAQPLLRPNAGSS"	Db	2431	ACCCATTGCGCTCCCAATCGGGCTAAAGGCTTGCCATTTGCTCTGATGGCTTAAGTGCC	2490
	43. .57	Qy	1081	TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTTCCATGGTTCTTCTCCATGAC	1140
	284. .1900	Db	2491	TGGGTTTCATCTTAATAGCTGAACACTAGTCACTGGGTTTCCATGGTTCTTCTCTGTGAC	2550
	/note="syncytin"	Qy	1141	CGAGGCTTCTAATAGAGCTATAAACAATCCATGCGCCCAAGATTCCCAATTCCTTGGTA	1200
	43. .57	Db	2551	CCACGCTTCTAATAGAGCTATAAACAATCCATGCGCCCAAGATTCCCAATTCCTT-GAA	2609
	284. .1900	Qy	1201	TCGTGAGGCAAGAACCCAGGTGAGAGAAAGTGGGCTTGGCCACCATTTGGGAAG	1257
	/product="envelope glycoprotein" /protein_id="AAQ17603.1" /db_xref="GI:33411093" /translation="MALPYHFLFTLLPSFTLTAPPCKMTSSSPYOEFLWRMORP GNIDARSLSKGTPTFTAHMPNCVHSATLQNHANTHYWTGWINPSPGVLGV HTVLVLFNTTLGLHEVDAQREKHVEVISQLTRVHGTSSPYKGLDLSKLHETLT LVGLVSNLEITHTSNLTCKVFSNTTNTNSQCIRWTPPTQIVCLPSGIFVCGTSA YRLGSSSSEMFSLPVPMTIYTDQDLYNYVISKPRNRPVILPFIAGVIGLALG TGIGGTTSTQFYKLSOELNGDMERVADSLVTDQLNSLAALAVLQNRALDLTAE RGTCIFLGECCYYNQSIGIVTEKVEIRDRIOREARELRNTPKGLLSQMPWILP FLGPLAAILLLFLGFCIFNLLNVFVSSRIEAVKLQMEPKMQSKTKIYRRPLDRPASP RSDVNDIKGTPPEEISAAQPLLRPNAGSS"	Db	2610	TCCATAGGCCAAGAACCCAGGTGAGAGAAAGTGGGCTTGGCCACCATTTGGGAG	2666
	43. .57	Qy	1201	TCGTGAGGCAAGAACCCAGGTGAGAGAAAGTGGGCTTGGCCACCATTTGGGAAG	1257
	284. .1900	Db	2610	TCCATAGGCCAAGAACCCAGGTGAGAGAAAGTGGGCTTGGCCACCATTTGGGAG	2666

AF52056352	AF52056352	2694 bp	DNA	linear	PRI 11-FEB-2004
LOCUS	Homo sapiens individual 71 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.				
DEFINITION					
ACCESSION	AF520564				
VERSION	AF520564.1	GI:33411096			
KEYWORDS					
SEGMENT					
SOURCE	2 of 2				
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 2694)				
TITLE	Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.				
JOURNAL	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology				
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)				
AUTHORS	14757826				
TITLE	2 (bases 1 to 2694)				
JOURNAL	Mallet,F., Bouton,O. and Oriol,G.				
FEATURES	Direct Submission				
source	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France				
misc_feature	Location/Qualifiers				
CDS	1..2694				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/isolate="individual 71 allele B"				
	/db_xref="taxon:9606"				
	/chromosome="7"				
	/map="7q21-q22"				
	/sex="male"				
	/cell_type="PBMC's"				
	/note="African				
	endogenous_virus: HERV-W"				
	43..57				
	/note="gp125 acceptor site"				
	284..1900				
	/note="syncytin"				
	/codon_start=1				
	/product="envelope glycoprotein"				
	/protein_id="AAQ17604.1"				
	/db_xref="GI:33411097"				
	/translation="MALPYHFLFTVLLPSFTLAPPRCMTSSSPYQBELWRKQRP GNIDAPSRSLSGTPTTAHMPNRCYHSATLCHMANTHWIGKMINPSCPLGLV TVCWTFYQTMDSGGVQDOAREKHVKEVISQLTRVHGTSSPYKGLDLKSLHETLR HTLVLSFNTLTGLHEVSAQNPTNCWICLPLNRPYVSIPEVQWNNFSEINTTSV LVGPLVSNLEITHSNLTCVFSNTVTNSQCIRWVTPPTQIVCLPSGIFPVGCTSA YRLNGSSSEMTCSFLVPPMTIITEODLYSVISKPRNRPVILPFVIGALVGLGAL TGIGGITTSQFYKLSQELNGDMERVADSLVTIQDOLNSLAUVLQNRALDILTAE RGGCLFUGESCCTVNGSGVITEKVEIRDRIORRAEELRNTGFWLLSQWMPWILP FLGPLAAIILLLLPFCPIFNLLVNFVFSRIEAVLQMEPRMQSKTIYRRLDRPASP RSDVNDIKGTPEEISAAQLLPNSAGSS"				
	1705..1720				
	/note="gp125 acceptor site"				
	1959..2694				
ORIGIN					
Query Match	77.9%; Score 1035.2; DB 9; Length 2694;				
Best Local Similarity	90.1%; Pred. No. 3e-304;				
Matches 1133; Conservative	0; Mismatches 111; Indels 13; Gaps 2;				
Qy	1 TCAAAATCGAAGCTTTAGACTTGCTAAACCGCCAAAGAGAGGGGAACTGTTATTTT 60				
Db	1423 TCAAAATCGAAGCTTTAGACTTGCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 1482				
Qy	61 AGGGGAAGAACTGCTTGTAGTATGTTAACTCTGAATCTACTGAGAAGTTAAAGA 120				
Db	1483 AGGGGAAGAACTGCTTGTATTTATTTATTAATCAATCCGAATCGTCACTGAGAAGTTAAGA 1542				
Qy	121 AATTGAGATCGAATATTAATGTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGCGCT 180				

```
Db      2610 TCCATAAGGCCAAGAACCCAGGTCAGAGAACACAGGAGCTTGCCACCATCTTTGGGAG 2666

RESULT 32
AX329572
LOCUS      56093 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION      Sequence 81 from Patent WO0194629.
ACCESSION      AX329572
VERSION      AX329572.1  GI:18102550
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
  Horrigan, S., Soppet, D.R. and Weaver, Z.
  Cancer gene determination and therapeutic screening using signature
  gene sets
JOURNAL      Patent: WO 0194629-A 81 13-DEC-2001;
              Avalon Pharmaceuticals (US)
FEATURES
source      1..56093
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"

ORIGIN
Query Match      77.9%; Score 1035.2; DB 6; Length 56093;
Best Local Similarity 90.1%; Pred. No. 3.2e-304;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;

QY      1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCAGGAGGGGAACTGTTATTTT 60
Db      37018 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCAGGAGGGGAACTGTTATTTT 37077

QY      61 AGGGGAAGAGTCTGTATGATCTTAATCAATCTGGAATCATTAATGAGAGGTTAAAGA 120
Db      37078 AGGGGAAGAGTCTGTATGATCTTAATCAATCTGGAATCATTAATGAGAGGTTAAAGA 37137

QY      121 AATTGAGATCGAATATAATGTAGACGAGGAGCTTCAAAACACTGCACCCCTGGGGCCT 180
Db      37138 AATTGAGATCGAATATAACAGTGTAGACGAGGAGCTTCAAAACACTGCACCCCTGGGGCCT 37197

QY      181 CCTCAGCAATGATGCGCCCTGGAGCTCTCCCTTTCTTAGACCTCTAGCAGCTATATATT 240
Db      37198 CCTCAGCAATGATGCGCCCTGGAGCTCTCCCTTTCTTAGACCTCTAGCAGCTATATATT 37257

QY      241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTTCTTGTAAAGTTTGTCTCTTCCAGAAAT 300
Db      37258 GCTACTCTCTTTGGACCCCTGTATCTTCAACTTCTTGTAAAGTTTGTCTCTTCCAGAAAT 37317

QY      301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCGATCCATGACTAA 360
Db      37318 CGAAGCTGTAAAGCTA-----CAATGGAGCCCAAGATGCGATCCCAAGACTAA 37365

QY      361 AATCTACCGTGGACCCCTGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db      37366 GATCTACCGGAGACCCCTGACCGGCTGTAGCCCAAGATGCTGATGTTAATGACATCAA 37425

QY      421 AGTCACCCCTCCGAGGAAATCTCACTGCACAAACCCCTACTACCTCCCAATTCAGTAGG 480
Db      37426 AGGCACCCCTCCGAGGAAATCTCAGTGCACAAACCCCTACTACCTCCCAATTCAGTAGG 37485

QY      481 AAGCAGTTAGACGAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
Db      37486 AAGCAGTTAGACGAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 37545

QY      541 GGGTGGACTGAGAGACGAGTACTGAGTTTCTTAGGCTGACTAAGAAATCCCNAGCCT 600
Db      37546 TGGGGGACTGAGAGACGAGTACTGAGTTTCTTAGGCTGACTAAGAAATCCCNAGCCT 37605

QY      601 ANCTGGGAAGGTGACCCGATCCATCTTTTAAACATGGGGCTTTGCAACTTAGCTCACACCCG 660

Db      37606 AGCTGGGAAGGTGACCCATCCACTTTTAAACACGGGGCTTTGCAACTTAGCTCACACCTG 37655
QY      661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAGAGGTAAGAGCAATAGCC 720
Db      37666 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAGAGGTAAGAGCAATAGCC 37725
QY      721 AATCATCTATTGCTGAGAGCAGACGCGGAAAGGAGGATTTGGATATATAAATCAGGCA 780
Db      37726 AATCATCTATTGCTGAGAGCAGACGCGGAGGAGCAATATGATCGGATATAAATCAGGTC 37785
QY      781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTT 840
Db      37786 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTT 37845
QY      841 CACTCTATTTCATCTCTATTAAATCATGCAACTGCACTCTCTCTGTCCTGTTTATGG 900
Db      37846 CATGCTATTTCATCTCTATTAAATCTTGCACTGCACTCTCTCTGTCCTGTTTATGG 37905
QY      901 CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGCTCAGACCCGCT 960
Db      37906 CTTGAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCAGCTCAGACCCGCT 37965
QY      961 GCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCACTGCTCTGATCCAGCGAGCT 1020
Db      37966 GCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCACTGCTCTGATCCAGCGAGCT 38025
QY      1021 ACCCAATTCGCACTCCCGATCAGGCTAAAGGCTTTGCAATTTCTCTGATGGCTAAGTGCC 1080
Db      38026 ACCCAATTCGCGCTCCCAATCGGCTTAAGGCTTTGCAATTTCTCTGATGGCTAAGTGCC 38085
QY      1081 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTCACTGGGTTCCATGTTCTCTTCCATGAC 1140
Db      38086 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTCACTGGGTTCCATGTTCTCTTCTGTGAC 38145
QY      1141 CCACGGCTTCTTAATAGAGCTAATCACTCACCGCATGGCCCAAGATTCCATCTCTTGGTA 1200
Db      38146 CCACAGCTTCTTAATAGAGCTAATCACTCACCGCATGGCCCAAGATTCCATCTCTT-GAA 38204
QY      1201 TCTGTGAGGCCAAGAACCCAGCTCAGAGAAAGTGGAGCTTGCCACCATTTGGGAAG 1257
Db      38205 TCCATAAGGCCAAGAACCCAGCTCAGAGAAAGTGGAGCTTGCCACCATTTGGGAG 38261

RESULT 33
HSAC000064      56093 bp      DNA      linear      PRI 13-NOV-1996
LOCUS      Human BAC clone RG083M05 from 7q21-7q22, complete sequence.
DEFINITION      HSAC000064
ACCESSION      AC000064
VERSION      AC000064.1  GI:1669369
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Pauley, A.
  The sequence of H. sapiens BAC clone RG083M05
  Unpublished (1996)
REFERENCE
2. (bases 1 to 56093)
  Waterston, R.
  Direct Submission
  Submitted (13-NOV-1996)
JOURNAL      Genome Sequencing Center
              Department of Genetics, Washington University
              St. Louis, MO 63108, USA
              e-mail: saplens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
sections once, or longer because we provide a small overlap between
neighboring submissions.
```

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION:

This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 978SK. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).
VECTOR: pBEO
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H_RG083M05; actual end is at 56093 of H_RG083M05

This clone contains STS sWSS1725.

FEATURES

source

1..56093
/location=Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-7q22"
/clone="H_RG083M05"
/clone.lib="CITB-978SK-B"
complement(838..1131)
/rpt_family="ALU"
c1360..16971

repeat_region

gene

/gene="WUGSC:H_RG083M05.1"
join(c1360..1503,4181..4370,4587..4774,6422..6556,9483..9547,11631..11773,11864..12021,13131..13296,14885..14988,16349..16546,16837..16971)
/gene="WUGSC:H_RG083M05.1"
/note="ATPase; strong similarity to peroxisome biosynthesis protein PASI (PID:g1172019); coded for by human cDNA C04279 (NID:g1467530)"
/codon_start=1
/protein_id="BAB46346.1"
/db_xref="GI:1669371"

/translation="KRLENIQKTEVAFSEAVWQPSVVLDDLDLGLPAVPEHEH
SPDAQRCEILCNVTKNKLDCDKINKFTDLDQHVAKETGGFVARDFTVLVDRAIHSRL
SRQSTIREKLVLTLDFQKALRGFLPASRLVNLKPRDLGWDKIGLHVEVRILMD
TIQLPAKCLKKEKYPFLANLPKIRQRTGILLYGPGTGKTLAAGVIAREGRMFVSV
KGPELLSKYIGASEQAVRDIFIRAQAAPCILFPDEFESIAPRRHGNTGVTDRVNVQ
LLTOLDGVEGLQGVYVLAATSRPDLIDPALLRPGRLDKVCVCPDQVITIGYLSKTC
QMLRSFLVSRLEILNVLSDSLPLADDVLDLQHVASVTDSTGTADLKALLYNQALELHG
MLLSKMSILPDESFPNRYLFGSSYESELNGTSSDLSSQCLASPSMTQDLPGVP
GKQQLFSPQVPLVRTASQSGQELTQEQRDQLRADISIIRKGYRSQSGEDEMNPQGP
KTRLAISQHLMTALGHTRPISIEDDKNFAL"
complement(4948..5130)
/rpt_family="ALU"
complement(6581..7133)
/rpt_family="L1"
complement(7767..8037)
/rpt_family="ALU"
complement(8186..8472)
/rpt_family="ALU"
8473..8625
/gene="WUGSC:H_RG083M05.1"
/note="match to human 3' EST H75782 (NID:g1049794), bases 287-444"

repeat_region

repeat_region

repeat_region

repeat_region

misc_feature

/note="match to human 3' EST H75782 (NID:g1049794), bases 287-444"
8841..9161
/gene="WUGSC:H_RG083M05.1"
/note="match to human 5' EST H75921 (NID:g1050050), bases 21-348"

misc_feature

misc_feature

9481..9547
/gene="WUGSC:H_RG083M05.1"
/note="match to human 5' EST N22627 (NID:g1130501), bases

276-343"
complement(12612..12907)
/rpt_family="ALU"
13670..13793
/gene="WUGSC:H_RG083M05.1"
/note="match to human 5' EST H41382 (NID:g17434), bases 143-266"
13794..13877
/rpt_family="ALU"
13878..13906
/gene="WUGSC:H_RG083M05.1"
/note="match to human 5' EST H41382 (NID:g917434), bases 30-58"
13907..14104
/rpt_family="ALU"
complement(14110..14137)
/rpt_family="L1"
complement(15618..15907)
/rpt_family="ALU"
17227..17522
/rpt_family="ALU"
18667..19235
/note="match to human fetal brain 5' EST D61494 (NID:g970409), bases 1-255, and to human 3' EST R07476 (NID:g759399)"
19550..19670
/rpt_family="ALU"
21507..37303
/note="similarity to various SS-RNA virus polypeptides; pseudogene; region of matches and close matches to multiple human ESTs, see R68740 (NID:g842257)"
37316..37489
/note="Grail prediction, score = 80"
/evidence=not_experimental
complement(38938..39224)
/rpt_family="ALU"
39225..39707
/note="match to multiple human ESTs, see N30113 (NID:g1148633)"
39800..40085
/rpt_family="ALU"
complement(40247..40538)
/rpt_family="ALU"
complement(40632..40924)
/rpt_family="ALU"
complement(42283..42891)
/rpt_family="ALU"
complement(45474..45613)
/rpt_family="ALU"
complement(45614..45737)
/note="match to human 3' EST H48898 (NID:g988738), bases 129-333"
complement(46107..47026)
/note="match to multiple human ESTs, see N81064 (NID:g1243765), H48897 (NID:g988737), and W78831 (NID:g273146)"
complement(47027..47318)
/rpt_family="ALU"
complement(47365..47782)
/note="match to multiple human ESTs, see W37495 (NID:g1319089)"
47898..48115
/note="match to human 5' EST H62306 (NID:g1015138), bases 93-368"
complement(48116..48405)
/rpt_family="ALU"
complement(48406..48584)
/note="match to human 3' EST N29952 (NID:g1148472), bases 290-455, and 5' EST R12730 (NID:g765806)"
complement(48787..49405)
/rpt_family="ALU"
complement(49406..49534)
/note="match to human 3' EST R65794 (NID:g838432), bases

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Direct Submission
 Submitted (02-OCT-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 5 (bases 1 to 149194)
 Waterston,R.H.
 Direct Submission
 Submitted (16-NOV-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 6 (bases 1 to 149194)
 Waterston,R.H.
 Direct Submission
 Submitted (03-JAN-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 7 (bases 1 to 149194)
 Waterston,R.H.
 Direct Submission
 Submitted (06-FEB-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 8 (bases 1 to 149194)
 Waterston,R.
 Direct Submission
 Submitted (01-MAR-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Nov 16, 2000 this sequence version replaced gi:4835815.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics

 Center project name: H_RG010G05

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 Clone CTB-10G5 is from the first release of the human BAC library CITB-97ASK-B. The library contains cloned DNA from the male fibroblast cell line 97ASK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
 VECTOR: pBelOBAC11
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP5-911H5, 200 base pair overlap. Actual start of this clone is at base position 195 of CTB-10G5; actual end is at base position 150532 of CTB-10G5.

FEATURES
source
 The clone CTB-10G5 contains the entire sequence of CTB-83M5.
 1. .149194
 /location="Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7q21-7q22"
 /clone="CTB-10G5"
 /clone_lib="CITB-97ASK-B"
 1. .1634
 /rpt_family="L1"
 2248..2388
 /note="match to EST BG752883 (NID:G14063536)"
 2248..2387
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 2248..2387
 /note="match to EST AW579261 (NID:G7254310)"
 2248..2387
 /note="match to EST BG766882 (NID:G14077535)"
 2248..2387
 /note="match to EST B1160365 (NID:G14620366)"
 2248..2374
 /note="similar to Homo sapiens EST BF758865 (NID:G12106765)"
 2248..2287
 /note="match to EST AV686676 (NID:G10288539)"
 2253..2387
 /note="match to EST AU123510 (NID:G10948226)"
 2344..2387
 /note="match to EST AV686676 (NID:G10288539)"
 2696..3066
 /rpt_family="MaLR"
 3108..3332
 /rpt_family="Alu"
 3540..3628
 /note="match to EST AA425526 (NID:G2106267) zw48b03.rl"
 3540..3628
 /note="similar to Mus musculus EST BB253526 (NID:G8946272)"
 3542..3632
 /note="match to EST AV686676 (NID:G10288539)"
 3542..3628
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 3542..3628
 /note="match to EST AU123510 (NID:G10948226)"
 3542..3628
 /note="match to EST BG766882 (NID:G14077535)"
 3542..3628
 /note="match to EST B1160365 (NID:G14620366)"
 3542..3609
 /note="match to EST AW579261 (NID:G7254310)"
 3562..3628
 /note="match to EST BE272564 (NID:G9146913)"
 3598..3628
 /note="match to EST BE299708 (NID:G9183456)"
 3714..3785
 /note="match to EST BG260659 (NID:G12770475)"
 3717..3785
 /note="similar to Mus musculus EST BE994936 (NID:G10678674)"
 3719..3785
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 3719..3785
 /note="match to EST AA425526 (NID:G2106267) zw48b03.rl"
 3719..3785
 /note="match to EST AU123510 (NID:G10948226)"
 3719..3785
 /note="match to EST BE272564 (NID:G9146913)"
 3719..3785
 /note="match to EST BE299708 (NID:G9183456)"
 3719..3785
 /note="match to EST BE299708 (NID:G9183456)"
 3719..3785

repeat_region
 1. .1634
 /rpt_family="L1"
 2248..2388
 /note="match to EST BG752883 (NID:G14063536)"
 2248..2387
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 2248..2387
 /note="match to EST AW579261 (NID:G7254310)"
 2248..2387
 /note="match to EST BG766882 (NID:G14077535)"
 2248..2387
 /note="match to EST B1160365 (NID:G14620366)"
 2248..2374
 /note="similar to Homo sapiens EST BF758865 (NID:G12106765)"
 2248..2287
 /note="match to EST AV686676 (NID:G10288539)"
 2253..2387
 /note="match to EST AU123510 (NID:G10948226)"
 2344..2387
 /note="match to EST AV686676 (NID:G10288539)"
 2696..3066
 /rpt_family="MaLR"
 3108..3332
 /rpt_family="Alu"
 3540..3628
 /note="match to EST AA425526 (NID:G2106267) zw48b03.rl"
 3540..3628
 /note="similar to Mus musculus EST BB253526 (NID:G8946272)"
 3542..3632
 /note="match to EST AV686676 (NID:G10288539)"
 3542..3628
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 3542..3628
 /note="match to EST AU123510 (NID:G10948226)"
 3542..3628
 /note="match to EST BG766882 (NID:G14077535)"
 3542..3628
 /note="match to EST B1160365 (NID:G14620366)"
 3542..3609
 /note="match to EST AW579261 (NID:G7254310)"
 3562..3628
 /note="match to EST BE272564 (NID:G9146913)"
 3598..3628
 /note="match to EST BE299708 (NID:G9183456)"
 3714..3785
 /note="match to EST BG260659 (NID:G12770475)"
 3717..3785
 /note="similar to Mus musculus EST BE994936 (NID:G10678674)"
 3719..3785
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 3719..3785
 /note="match to EST AA425526 (NID:G2106267) zw48b03.rl"
 3719..3785
 /note="match to EST AU123510 (NID:G10948226)"
 3719..3785
 /note="match to EST BE272564 (NID:G9146913)"
 3719..3785
 /note="match to EST BE299708 (NID:G9183456)"
 3719..3785

misc_feature
 1. .1634
 /rpt_family="L1"
 2248..2388
 /note="match to EST BG752883 (NID:G14063536)"
 2248..2387
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 2248..2387
 /note="match to EST AW579261 (NID:G7254310)"
 2248..2387
 /note="match to EST BG766882 (NID:G14077535)"
 2248..2387
 /note="match to EST B1160365 (NID:G14620366)"
 2248..2374
 /note="similar to Homo sapiens EST BF758865 (NID:G12106765)"
 2248..2287
 /note="match to EST AV686676 (NID:G10288539)"
 2253..2387
 /note="match to EST AU123510 (NID:G10948226)"
 2344..2387
 /note="match to EST AV686676 (NID:G10288539)"
 2696..3066
 /rpt_family="MaLR"
 3108..3332
 /rpt_family="Alu"
 3540..3628
 /note="match to EST AA425526 (NID:G2106267) zw48b03.rl"
 3540..3628
 /note="similar to Mus musculus EST BB253526 (NID:G8946272)"
 3542..3632
 /note="match to EST AV686676 (NID:G10288539)"
 3542..3628
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 3542..3628
 /note="match to EST AU123510 (NID:G10948226)"
 3542..3628
 /note="match to EST BG766882 (NID:G14077535)"
 3542..3628
 /note="match to EST B1160365 (NID:G14620366)"
 3542..3609
 /note="match to EST AW579261 (NID:G7254310)"
 3562..3628
 /note="match to EST BE272564 (NID:G9146913)"
 3598..3628
 /note="match to EST BE299708 (NID:G9183456)"
 3714..3785
 /note="match to EST BG260659 (NID:G12770475)"
 3717..3785
 /note="similar to Mus musculus EST BE994936 (NID:G10678674)"
 3719..3785
 /note="match to EST AA149693 (NID:G1720635) zn99d06.rl"
 3719..3785
 /note="match to EST AA425526 (NID:G2106267) zw48b03.rl"
 3719..3785
 /note="match to EST AU123510 (NID:G10948226)"
 3719..3785
 /note="match to EST BE272564 (NID:G9146913)"
 3719..3785
 /note="match to EST BE299708 (NID:G9183456)"
 3719..3785

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

FEATURES		d'Italie, Lyon 69364 cedex 07, France
source		Location/Qualifiers
		1..2694
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/isolate="individual 22 allele B"
		/db_xref="taxon:9606"
		/chromosomes="7"
		/map="7q21-q22"
		/sex="male"
		/cell_type="PBMC's"
		/note="Caucasian"
		endogenous_virus: HERV-W"
misc_feature		43..57
		/note="splice acceptor site"
CDS		284..1900
		/note="syncytin"
		/codon_start=1
		/product="envelope glycoprotein"
		/protein_id="AAQ17568.1"
		/db_xref="GI:33410953"
		/translation="MALPYHFLFTVLLPSFTLTAPPCCRCMTSSPYQEFILWRQRP GNIDAPSYSLGKGTPTFTATHPRNCVHSATLCMHANTHYWTGMINPSPCGGLGV TVCWYFTQTMGSDGGVODQAREKHVEVISQLTOVHGTSPPYKGLDLKSLHEITLT HRLVSLFNTLTGLHEVSAQNTNCWICLPINFRPVVSI PVPEOWNFESTREINTSV LVGPLVNLIEITHISLNTCVKFSNTYTNQCIKRWTPPTQIVCLPSGIPFVCGTSA YCLNGSSSMCFSLSPVPMPTIYEODLYNYVISKPRNKRVPILFPVIGAGVIGALG TGIGITTSQFYKLSBELNGDMERVADSLVTLQDPLNSLAALVONRRALDILLAE RGCTCLFLGEECYVYNQSGIVTEKVEIRDIQRABELRNTGPMGLLSQWMPWILP FLGPLAAIILLLLFGPICFNLLVNFVSSRIEAVKLQMEPKMQSKTIYRRPLDRASP RSDVNDIKGTPPEEISAAQPLLRPNSSGS"
misc_feature		1705..1720
		/note="splice acceptor site"
LTR		1959..2694
ORIGIN		
Query Match		77.8%; Score 1033.6; DB 9; Length 2694;
Best Local Similarity		90.1%; Pred. No. 9.2e-304;
Matches 1132; Conservative		0; Mismatches 112; Indels 13; Gaps 2;
QY	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTTATTTT 60
DB	1423	TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGGAACCTGTTTATTTT 1482
QY	61	AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCAATTAAGAGAAAGTTAAAGA 120
DB	1483	AGGGGAAGAAATGCTGTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542
QY	121	AATTTGAGATCGAATATATGTAAGCAGAGGACCTTCAAAACACTGCACCTCGGGGCT 180
DB	1543	AATTCGAGATCGAATACAGCTAGAGCAGAGGAGCTTCGAACACTGGACCTCGGGGCT 1602
QY	181	CCTCAGCAATGAGTGCCTGGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
DB	1603	CCTCAGCAATGAGTGCCTGGATTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 1662
QY	241	TTTACTCTCTTTGGACCTGTAFCTTCAACTCTCTTTGTTAAAGTTGTCTCTCCAGAA 300
DB	1663	GCTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTGTTAAAGTTGTCTCTCTCCAGAA 1722
QY	301	TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA 360
DB	1723	CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATGAGTCCCAAGACTAA 1770
QY	361	AATCTACCTGGACCTCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
DB	1771	GATCTACCGAGACCTCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 1830
QY	421	AGTCACCCCTCCGAGAGAAATCTCAACTGCACAAACCCCTACTACCTCCAAATTCAGTAG 480
DB	1831	AGGCACCCCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
QY	481	AAGCAGTTAGGAGCTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
DB	1891	AAGCAGTTAGGAGCTGTCATCGCCAACTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGA 1950
QY	541	GGTGGAGCTAGAGACAGAGCTAGCTGATTTCTTAGGCTGACTAAGAAATCCCNAGCCT 600
DB	1951	TGGGGGAGCTAGAGACAGAGCTAGCTGATTTCTTAGGCTGACTAAGAAATCCCNAGCCT 2010
QY	601	ANCTGGGAAGGTGACCGATCCATCTTTAAACATGGGGCTTGCAACTTAGCTTCACACCCG 660
DB	2011	AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGCTTCACACCTG 2070
QY	661	ACCAATCAGAGAGCTCATTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
DB	2071	ACCAATCAGAGAGCTCATTAAATGCTTAATAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
QY	721	AATCATCTATTCGCTGAGAGCAGCGGGAAGGAGCAAGGATGGGATATAAACTCAGGCA 780
DB	2131	AATCATCTATTCGCTGAGAGCAGCGAGGAGGCAATGATCGGGATATAAAACCCAGTC 2190
QY	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTT 840
DB	2191	TTGAGCCGGAACCGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGCTCTGTTTT 2250
QY	841	CACCTCTATTTCACCTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTATGG 900
DB	2251	CATGCTATTTCACCTCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTACCG 2310
QY	901	CTCAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGTCACAGACCCGCT 960
DB	2311	CTTGAGCTGAGCTTTGCTCGCCATCCACACTGCTGTTTGGCCACCGTCACAGACCCGCT 2370
QY	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCAGGT 1020
DB	2371	GCTGACTTCCATCCCTCTCGATCATGAGGGTGTCCGCTGTGCTCTGATCCAGCAGGC 2430
QY	1021	ACCAATTCACCTCCCGATCAGGCTAAAGGCTTGCATTTCTCTGATCGCTAAGTGCC 1080
DB	2431	ACCAATTCGCGCTCCCAATTTGGCTAAAGGCTTGCATTTCTCTGATCGCTAAGTGCC 2490
QY	1081	TGGTTTGTCTTAATAGAACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB	2491	TGGTTTGTCTTAATAGAACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550
QY	1141	CAAGGCTTCTAATAGAGCTATAAACAACCTCAGCAGTGGCCCAAGATTTCATTCCTTGGTA 1200
DB	2551	CAAGGCTTCTAATAGAGCTATAAACAACCTCAGCAGTGGCCCAAGATTTCATTCCTT-GAA 2609
QY	1201	TCTGTGAGCCCAAGAACCCCGATCAGAAAGTGAAGGCTTGCACCATTTGGGAAG 1257
DB	2610	TCCATAAGCCCAAGAACCCCGATCAGAAACAGAGGCTTGCACCATTTGGGAG 2666
RESULT 37		
AF520499S2		2694 bp DNA linear PRI 11-FEB-2004
LOCUS		Homo sapiens individual 30 allele B, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.
DEFINITION		AF520500
ACCESSION		AF520500.1 GI:33410968
VERSION		2 of 2
KEYWORDS		Homo sapiens (human)
SEGMENT		Homo sapiens
SOURCE		Homo sapiens
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE		1 (bases 1 to 2694)
AUTHORS		Bonlaet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Mallet, B., Lucotte, G., Duret, L., and Mandrand, B.
TITLE		The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
PUBMED		14757826
REFERENCE		2 (bases 1 to 2694)

AUTHORS	Mallet, F., Bouton, O. and Oriol, G.
TITLE	Direct Submission
JOURNAL	Submitted (07-JUN-2002) Retrovirology Department, UMR 2142 CNRS-bioMarieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France
FEATURES	Location/Qualifiers
source	1..2694
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/isolate="individual 30 allele B"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7q21-q22"
	/sex="female"
	/cell_type="PBMC's"
	/note="Caucasian"
	endogenous_virus: HERV-W"
misc_feature	43..57
	/note="splice acceptor site"
CDS	284..1900
	/note="syncytin"
	/codon_start=1
	/product="envelope glycoprotein"
	/protein_id="AAQ17572.1"
	/db_xref="GI:33410969"
	/translation="MALPYHIFLFTVLLPSELTAPPCCRCKTSSPYOEFLEWRQRP GNIDPYSRUSKSGTPTFAHMPRNCHTSATLGMHANTHYWTKMINPSCPGLGV TVCMYFTQMGDDGGVDQAKHKEVIESQLTRVHGTSPPYKLDLSKLHETLRT HTRVLSENLITLGLHEVSAQNPTNCICLPLNFRPVYSIPVQMNWFSTEINTSV LKGPLVSNLEITHTSNLCVRFSTNTYTTNSQCIRWVPTQIVCLPSGIFFVCGTSA YRLGNSSSECFLSFLVPMTIYTEQDLYSVISKPRNKRVPILPFVIGAGVLGALG TGIGITSTQFYKLSQELNGDMERVADSLVTLQDOLNSLAADVQLNRRALDILTAE RGKTCFLGEECCYVNGSGIVTEYKZIRIRIQRABELRNTGFWGLLSQWMPKILP FLGLAAAILILLFGPICFNLNLNVFVSRIEAVLQMEPKMQSKTKIYRREDRPASP RSDVNDIKGTPEEISAAQPLLRNSAGS"
misc_feature	1705..1720
	/note="splice acceptor site"
LTR	1959..2694
ORIGIN	
Query Match	77.8%; Score 1033.6; DB 9; Length 2694;
Best Local Similarity	90.1%; Pred. No. 9.2e-304;
Matches 1132; Conservative	0; Mismatches 112; Indels 13; Gaps 2;
QY	1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAGAGAGGGGGAACCTGTTTATTTT 60
Db	1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCTGAAAGAGGGGGAACCTGTTTATTTT 1482
QY	61 AGGGGAAGAAATGCTGTATGTTTAAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA 120
Db	1483 AGGGGAAGAAATGCTGTATGTTTAAATCAATCCGGNATCGTCACTGAGAAAGTTAAAGA 1542
QY	121 AATTGAGATCGAATATAATAGTAGACGAGAGACCTTCAAACACTGCAACCTGGGGCCT 180
Db	1543 AATTGAGATCGAATACAACGATAGACGAGAGAGCTTCAAAACACTGGGACCTGGGGCCT 1602
QY	181 CCTCAGCAATGATGACCTGGACCTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
Db	1603 CCTCAGCAATGATGACCTGGACCTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 1662
QY	241 TTCTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCCAGAA 300
Db	1663 GCTACTCTCTTTGGACCTGTATCTTAACTCTCTTGTAACTTTGTCTCTTCCAGAA 1722
QY	301 TGAAGCTGTAAGCTTCAAAATAGTTCTTCAATGGAAACCCGAGATGCACTCCATCTAA 360
Db	1723 CGAAGCTGTAAGCTTCAAAATAGTTCTTCAATGGAAACCCGAGATGCACTCCATCTAA 1770
QY	361 AATCTACCTGGACCTCTGGACCGCTCTGACATGCTCTGATGTTAAATGACATTTGA 420
Db	1771 GATCTACCGGACCTCTGGACCGCTCTGATGTTAAATGACATTTGA 1830
QY	421 AGTCACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTACTACCTCCAAATTCAGTAGG 480

Db	1831 AGGCACCCCTCTCTGAGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
QY	481 AAGCAGTTAGAGAGAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTTCTGTTGAGA 540
Db	1891 AAGCAGTTAGAGAGAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTTCTGTTGAGA 1950
QY	541 GGGTGACCTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db	1951 TGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 2010
QY	601 ANCTGGGAAGGTGACCCGATCTCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660
Db	2011 AGCTGGGAAGGTGACCCGATCTCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCTG 2070
QY	661 ACCAATCAGAGAGCTCACTAAAATGCTNAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 720
Db	2071 ACCAATCAGAGAGCTCACTAAAATGCTNAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 2130
QY	721 AATCATCTATTGCTGAGAGACAGGGGAAGGACAAAGGATTTGGGATATAAACTCAGGCA 780
Db	2131 AATCATCTATTGCTGAGAGACAGGGGAAGGACAAAGGATTTGGGATATAAACTCAGGCA 2190
QY	781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTGTATGGGAGCTCTGTTT 840
Db	2191 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTGTATGGGAGCTCTGTTT 2250
QY	841 CACTCTATTCTCACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGCTGTTTATGG 900
Db	2251 CATGCTATTCTCACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGCTGTTTATGG 2310
QY	901 CTCAAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTGTGGCCACCTCTCAGACCCGCT 960
Db	2311 CTTCAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTGTGGCCACCTCTCAGACCCGCT 2370
QY	961 GCTGACTTCACTCTTGGATCCAGCAGAGTGTCCACTGCTGCTCTGATCCAGCGAGGT 1020
Db	2371 GCTGACTTCACTCTTGGATCCAGCAGAGTGTCCACTGCTGCTCTGATCCAGCGAGGT 2430
QY	1021 ACCCATTTGCCACTCCCGATCAGGCTAAAGGCTTGCCATTTGCTCGATGGCTAAAGTGC 1080
Db	2431 ACCCATTTGCCCTCCCGATCAGGCTAAAGGCTTGCCATTTGCTCGATGGCTAAAGTGC 2490
QY	1081 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db	2491 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550
QY	1141 CCAGCGCTTCTTAATAGAGCTTATACACTCACCGCATGCCCAAGATTTCCATTTCTTGGTA 1200
Db	2551 CCAGCGCTTCTTAATAGAGCTTATACACTCACCGCATGCCCAAGATTTCCATTTCTTGGTA 2609
QY	1201 TCTGTGAGGCCCAAGAACCCGAGTCCAGAGAAAGTGGAGCTTGCCACCATTTGGGAA 1257
Db	2610 TCCATAGGCCCAAGAACCCGAGTCCAGAGAAAGTGGAGCTTGCCACCATTTGGGAG 2666

RESULT 38

AF520501S2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AF520501S2

Homo sapiens individual 34 allele A, envelope glycoprotein gene, complete cds, and 3' long terminal repeat, complete sequence.

AF520502

AF520502.1

GI:33410972

2 of 2

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2694)

Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.

The endogenous retroviral locus ERVWE1 is a bona fide gene involved

in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
14757826
2 (bases 1 to 2694)
Mallet, F., Bouton, O. and Oriol, G.
Direct Submission
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
Location/Qualifiers
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolates="individual 34 allele A"
/db_xref="taxon:9606"
/chromosomes="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
43..57
/notes="splice acceptor site"
284..1900
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAQ17573.1"
/db_xref="GI:33410973"
/translation="MALPYHIFLTVLLPSFTLTAPPCCRMWSSPSYQEFLEWRMORP
GNIDAPYSLSKGTFTFTAHMPCNHSATLCMHANTHYMTGMINPSPGGLGV
TWCTVFTQMGSDGVDQKREHKEVISOLTRVHGTSYKGLDLKSLHETLR
HPLVSLFNTLTGLHEVSAONPTNCWICLPLNFRPVSVIPVPEOWNFSEINTSV
LVGLVSNLEIHTISNLTCKFSNTYTNQCIKRWTPPTQIVCLPSIFVCGTSA
YCLNGSSMCPCLSLFVPMWITLYEQDLYSVISKPRNKRVPILFVIGAGVGLG
TGIGTITTSQFYKLSQELNGMDERVADSLVLODLSLAFLVQLNRALDLITAE
RGCTCLFLGECVYVQNGIVTEKVEIKRDIQRAEELRNTGPMGLLSQWMPWILP
FLGLPAAIILLFLGFCIFNLVNFVSSRIEAVKLQWEPKMQSKTKIYRPLDRPASP
RSDVNDIKGTPEPPEISNAQPLLRPSAGSS"
1705..1720
/notes="splice acceptor site"
1959..2694

misc_feature
CDS

Query Match 77.8%; Score 1033.6; DB 9; Length 2694;
Best Local Similarity 90.1%; Pred. No. 9.2e-304;
Matches 1132; Conservative 0; Mismatches 112; Indels 13; Gaps 2;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAAACCTGTTATTTT 60
DB 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGAAACCTGTTATTTT 1482

QY 61 AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCATTAAGAGAAAGTTAAAGA 120
DB 1483 AGGGGAAGAAATGCTGTTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542

QY 121 AATTGAGATCGAATATAATGATAGACAGAGGACCTTCAAAACACTCCACCTGGGCGCT 180
DB 1543 AATTGAGATCGAATATAATGATAGACAGAGGAGCTTCGAAACACTCGGACCTGGGCGCT 1602

QY 181 CCTCAGCAATGGATGCGCTCGGACTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT 240
DB 1603 CCTCAGCAATGGATGCGCTCGGACTCTCCCTCTCTTAGGACCTCTAGCAGCTATAATATT 1662

QY 241 TTTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTTGAAGTTGTCCTCTCCGAAT 300
DB 1663 GCTACTCTCTTTGGACCTGTATCTTTAACTCTCTTTGTTAACTTTGCTCTCTCCGAAT 1722

QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAATGGAAACCCAGATGCGAGTCCAGCTAA 360
DB 1723 CGAAGCTGTAAAGCTA-----CAATGGAGCCCAAGATGCGAGTCCAGCTAA 1770

QY 361 AATCTACCGTGGACCTCGGACCGCTGCTAGACTATGCTCTGATCTTAATGACATTGA 420

DB 1771 GATCTACCGCAGACCCCTGGACCGGCTGTAGTCCCAAGATCTGATGTTAATGACATCAA 1830
QY 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCAACACCCCTACTACTCACTCAATTCAGTAGG 480
DB 1831 AGGCACCCCTCTCTGAGGAAATCTCACTAGTGCACAACTCTACTAGCGCCCAATTCAGCAG 1890
QY 481 AAGCAGTTAGACAGTTGTACGCCAACCTCCCAACAGTACTTGGTGTTCCTCTGTTGAGA 540
DB 1891 AAGCAGTTAGACAGTTGTGCGGCCAACCTCCCAACAGTACTTGGTGTTCCTCTGTTGAGA 1950
QY 541 GGTGTGACTGAGACAGAGTACTAGTGGATTTCTTAGGCTGACTAAGATTCCTCCNAAAGCCT 600
DB 1951 TGGGGGACTGAGACAGAGTACTAGTGGATTTCTTAGGCTGACTAAGATTCCTCCNAAAGCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGCGGCTTGCAACTTACTGCTCACACCGG 660
DB 2011 AGCTGGGAAGGTGACCGCATCCATCTTTAAACATGCGGCTTGCAACTTACTGCTCACACCGG 2070
QY 661 ACCAATCAGAGAGTCTATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
DB 2071 ACCAATCAGAGAGTCTATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2130
QY 721 AATCATCTATTCCTGAGAGCAGCGGGAAGSACAAGGATTTGGGATATATAAATCTCAGGCA 780
DB 2131 AATCATCTATTCCTGAGAGCAGCGGGAAGSACAAGGATTTGGGATATATAAATCTCAGGCA 2190
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 840
DB 2191 TTGAGCCGGAACCGCAACCCCTTTGGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 2250
QY 841 CACTCTATTTCACTCTATTAAATCATGCAACTGCACCTCTTCTGGTCCGTTGTTTTATGG 900
DB 2251 CATGCTATTTCACTCTATTAAATCTTGAACCTGCACCTCTTCTGGTCCGTTGTTTTATGG 2310
QY 901 CTCAGCTAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCCGTGCACACCGCTCACACCGCT 960
DB 2311 CTTGAGCTAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCCGTGCACACCGCTCACACCGCT 2370
QY 961 GCTGACTCTCCATCCCTTTGGATCCAGAGTGTGCACCTGTCTCTGTATCCAGCGAGGT 1020
DB 2371 GCTGACTCTCCATCCCTTTGGATCCAGAGTGTGCACCTGTCTCTGTATCCAGCGAGGT 2430
QY 1021 ACCAATTCACCTCCCGATCAGGCTAAAGCTTGGCATTTGTTCTTCTGATGGCTAAAGTGC 1080
DB 2431 ACCAATTCACCTCCCGATCAGGCTAAAGCTTGGCATTTGTTCTTCTGATGGCTAAAGTGC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 2491 TGGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2550
QY 1141 CCACGGCTTCTAATAGAGTATAAATCACTACCGCATGGCCCAAGATTCCTTCTGCTGTA 1200
DB 2551 CCACGGCTTCTAATAGAGTATAAATCACTACCGCATGGCCCAAGATTCCTTCTGCTGTA 2609
QY 1201 TCTGTGAGGCAAGAACCCCGATCAGAGAAAGTGGGCTTGGCCACCATTTGGGAAG 1257
DB 2610 TCCATAAGGCCAAGAACCCCGATCAGAGAAAGTGGGCTTGGCCACCATTTGGGAAG 2666

RESULT 39
AF520507S2
LOCUS
DEFINITION
AF520507S2
complete cds, and 3' long terminal repeat, complete sequence.
AF520508
AF520508.1 GI:33410984
2 of 2
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2694)
REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2694)
Mallet, F., Bouton, O., Prudhomme, S., Cheynet, V., Oriol, G.,
Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.
The endogenous retroviral locus ERVWE1 is a bona fide gene involved
in hominoid placental physiology
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)
JOURNAL
PUBMED
14757826
2 (bases 1 to 2694)
Mallet, F., Bouton, O. and Oriol, G.
Direct Submission
Submitted (07-JUN-2002) Retrovirology Department, UMR 2142
CNRS-bioMerieux, Ecole Normale Supérieure de Lyon - 46 allée
d'Italie, Lyon 69364 cedex 07, France
FEATURES
source
1. .2694
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="individual 45 allele B"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21-q22"
/sex="male"
/cell_type="PBMC's"
/note="endogenous_virus: HERV-W"
misc_feature
43..57
/notes="splice acceptor site"
CDS
284..1900
/notes="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AA017578.1"
/db_xref="GI:33410993"
/translation="MALPHYFLFTVLLPSFTLTAPPCKRMTSSPYDEFLWRMRP
GNIDAPYSRLSGTFTFAHTMPNRYHSATLCMHANTHYWTGMINPSPGGILV
TICWTYFTQMGSDGGVDOAREKHVEVISOLTRVHGTSPPYKGLDLSKLHETLT
HPELVSLNTLTGLHEVSAONPTNCWICLPLNFRPVSI PVPEONNFSSTEINTSV
LVGLPVSNLEIHTSNLTGVKFSNTTYTNSQIRWVPTQIVCLPSGIFVCGTSA
YRCLNGSSMSFLVLPWMTIYTEQDLISYVISKPRNKRVPILFVIGAGVLGALG
TGIGTITTFQFYLSQELNGDMERVADSLVTLQDLSLAVALVIONRRALLTAE
RGCTFLAGECCYVQNSGIVTEKVEIRLQRAEELRNTGPGWLLSOMWPTLIP
FLGLAAIILLFLFGPCIFNLVNFVSSRIEAVKLQWPKVSKTKIYRPLDRPASP
RSDVNDIKGTPPEISNAQPLLRPNAGSS"
misc_feature
1705..1720
/notes="splice acceptor site"
LTR
1959..2694
ORIGIN
Query Match 77.88; Score 1033.6; DB 9; Length 2694;
Best Local Similarity 90.14; Pred. No. 9.2e-304;
Matches 1132; Conservative 0; Mismatches 112; Indels 13; Gaps 2;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGGAACCTGTTATTTT 60
DB 1423 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGGGGGAACCTGTTATTTT 1482
QY 61 AGGGGAAGAATGCTGTTAGTATGTTTAACTCAATCTGGAATCATTAAGAGAAAGTTAAAGA 120
DB 1483 AGGGGAAGAATGCTGTTATGTTTAACTCAATCGGAATCGTCACTGAGAAAGTTAAAGA 1542
QY 121 AATTGATGATGATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 1543 AATTGATGATGATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1602
QY 181 CCTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 1603 CCTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1662
QY 241 TTTACTCTCTTTGGACCCCTGATCTTCACTCTGTTTAAAGTTTGTCTTCTCCAGAAAT 300
DB 1663 GCTACTCTCTTTGGACCCCTGATCTTTAACTCTCTGTTTAACTCTTGTCTCTCCAGAAAT 1722

Search completed: February 21, 2005, 11:52:42
Job time : 5718.48 secs

QY 301 TGAAGCTGTAAAGCTACAATAAGTTCTTCAAATGGAAACCCAGATGCACTGATCCATGACTAA 360
DB 1723 CGAAGCTGTAAAGCTA-----CAAATGGAGCCCAAGATCAGTCCCAAGACTAA 1770
QY 361 AATCTACCGTGAACCCCTGGACCGGCTGCTAGATCTATCTCTGATGTTTAAATGACATTTGA 420
DB 1771 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACCATCTGATGTTTAAATGACATCAA 1830
QY 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG 480
DB 1831 AGGCACCCCTCTCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 1890
QY 481 AAGCAGTTAGACAGTTGTCAGCCAACTCCCAACAGACTTGGGTTTTCTCTGTTGAGA 540
DB 1891 AAGCAGTTAGACAGTTGTCAGCCAACTCCCAACAGACTTGGGTTTTCTCTGTTGAGA 1950
QY 541 GGGTGGCTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAGAAATCCCAAGCCCT 600
DB 1951 TGGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAGAAATCCCTAAGCCCT 2010
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCCG 660
DB 2011 AGCTGGGAAGGTGACCAATCCATCTTTAAACAGGGGCTTGCACCTTAGCTCACACCTG 2070
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGAGGTTAAGCAATAGCC 720
DB 2071 ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGAGGTTAAGCAATAGCC 2130
QY 721 AATCATCTATTGCTGAGACAGCGGGAAGAGAGGATTTGGGATATAAATCAGGCA 780
DB 2131 AATCATCTATTGCTGAGAGCAGCAGGAGGACAAATGATCGGGATATAAATCAGGCA 2190
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
DB 2191 TTGAGCCGCGCAACGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTT 2250
QY 841 CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCTGTTTTATGG 900
DB 2251 CATGCTATTTCATCTATTAAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGG 2310
QY 901 CTCAGCTAGCTTTTGTGGCATCCACACTGCTGTTTGGCCACCTGACAGACCCGCT 960
DB 2311 CTTGAGCTGAGCTTTTCCGCTCGCCATCCACACTGCTGTTTGGCCACCTGACAGACCCGCT 2370
QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCGAGG 1020
DB 2371 GCTGACTCCATCCCTCTGGATCATCAGGGTGTCCGCTGTCTCTGATCCAGCGAGG 2430
QY 1021 ACCCATTTGCCACTTCCCGATCAGGCTAAAGCTTTGCCATTTCTCTGATGGCTAAAGTGC 1080
DB 2431 ACCCATTTGCCCTTCCCAATCGGCTAAAGCTTTGCCATTTCTCTGATGGCTAAAGTGC 2490
QY 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCTACTGGTTCCATGTTCTCTTCCATGAC 1140
DB 2491 TGGGTTTGTCTTAATAGAACTGAACACTGGTCTACTGGTTCCATGTTCTCTTCTGTGAC 2550
QY 1141 CCACGGCTTCTAATAGAGCTATAAACAACCTACCGCATGGCCCAAGATTTCCATTTCTCTG 1200
DB 2551 CCACGGCTTCTAATAGAGCTATAAACAACCTACCGCATGGCCCAAGATTTCCATTTCTT 2609
QY 1201 TCTGTGAGGCCAAGAACCCAGGTCAGAGAAAGTGGGTTGGCCACCATTTGGGAAG 1257
DB 2610 TCCATAAGGCCAAGAACCCAGGTCAGAGAAACAGAGGCTTGGCCACCATTTGGGAG 2666

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:25:25 ; Search time 685.525 Seconds
(without alignments)
11476.370 Million cell updates/sec

Title: US-09-319-156b-12
Perfect score: 1329
Sequence: 1 tcaaatgaagaggttag.....gcaaggatccccagtaaca 1329

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues 8780412
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	99.8	1329	2	AAV43219 Multiple
2	1326	99.8	1329	2	Aax29704 Clone 5M6
3	1326	99.8	1329	12	ADG14849
4	1299	97.7	1329	9	ADB84403 MSRV asso
5	1035.2	77.9	56093	6	ADL61744 Colon ade
6	1029.4	77.5	9502	10	ADFS9718 Human con
7	1023.2	77.0	10499	3	ABN97929 Human ret
8	801.6	60.3	7466	5	Aas68626 DNA encod
9	798.2	60.1	21646	11	ACN4454 Human gen
10	778.8	58.6	1393	3	Aas31002 Human dia
11	775.4	58.3	2030	3	AAG63826 Nucleotid
12	774.2	58.3	7974	11	ACN88786 Breat ca
13	766.8	57.7	2074	6	AAD41225 Human EMB
14	747.6	56.3	2046	8	ACC46747 Human dit
15	731.8	55.1	2946	2	Aax77526 Human sec
16	731.8	55.1	2946	3	Aaz59468 Human sec
17	731.8	55.1	2946	10	ADC38776 Human con
18	730.8	55.0	2930	6	AAD24195 Human syn
19	730.2	54.9	2781	5	Aaf55630 Nucleotid
20	728.6	54.8	6394	5	Aas84210 DNA encod

21	722.6	54.4	7582	2	AAX25665 Complete
22	722.6	54.4	7582	3	Aax59215 Human end
23	721.2	54.3	1136	2	Aax25660 Human end
24	721.2	54.3	1136	3	Aax59210 3' pol ge
25	712.6	53.6	2782	2	Aax25661 Human end
26	712.6	53.6	2782	3	Aax59211 5' non co
27	712.6	53.6	2782	5	Aah20069 HERV-W en
28	704.8	53.0	1894	4	Aai14608 Probe #45
29	704.8	53.0	1894	4	Abas6337 Human foe
30	704.8	53.0	1894	4	Abas5980 Probe #46
31	704.8	53.0	1894	4	Abas5822 Human bre
32	704.8	53.0	1894	4	Abas5978 Probe #44
33	704.8	53.0	1894	4	Abas30018 Human bon
34	704.8	53.0	1894	4	Aak04516 Human bra
35	704.8	53.0	1894	4	Abas29670 Human liv
36	704.8	53.0	1894	5	Aai104422 Human gen
37	704.8	53.0	1894	6	Abas04589 Human gen
38	680.6	51.2	2942	6	Aas77313 DNA encod
39	651	49.0	893	5	Aas65964 DNA encod
40	643.4	48.4	46340	3	ABN97978 Human ret
41	643.4	48.4	161334	11	ACN44334 Human gen
42	639.8	48.1	783	2	Aax25669 Human end
43	637.8	48.0	780	7	ADS31352 Human gen
44	610.4	45.9	3464	10	ADE09587 Novel DNA
45	607	45.7	2784	3	ABN97930 Human ret

ALIGNMENTS

RESULT 1
AAV43219
ID AAV43219 standard; cDNA; 1329 BP.
XX
AC AAV43219;
XX
DT 29-DEC-1998 (first entry)
XX
DE Multiple sclerosis associated retrovirus fragment 6.
XX
XX Multiple sclerosis associated retrovirus; MSRV; MS; pol gene; gag gene;
KW env gene; rheumatoid arthritis-associated virus; ss.
XX
OS Multiple sclerosis associated retrovirus.
XX
FH Key Location/Qualifiers
CDS 2..490
FT /*tag= a
FT /product= "Encodes protein AAV71069"
FT /transl_except= (pos:77-79, appears to code for a stop
FT codon)
FT /transl_except= (pos:125-127, appears to code for a stop
FT codon)
FT /transl_except= (pos:137-139, appears to code for a stop
FT codon)
XX
XX WO9823755-A1.
XX
XX 04-JUN-1998.
XX
XX 26-NOV-1997; 97WO-IB001482.
XX
XX 26-NOV-1996; 96US-00756429.
XX
XX (INMR) BIO MERIEUX.
XX
XX Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
XX Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B;
XX WPI; 1998-322732/28.
XX P-PSDB; AAW71069.
XX
XX New nucleic acid from retroviruses - useful for diagnosis, prevention and

PS
XX
CC This sequence represents clone 5M6 from a novel multiple sclerosis
CC related virus type 1 (MSRV1). The sequence can be used in diagnostic,
CC prophylactic or therapeutic compositions to inhibit expression of a
CC multiple sclerosis related virus and/or virus associated with rheumatoid
CC polyarthritis. (Updated on 27-AUG-2003 to correct OS field.) (Updated on
CC 17-OCT-2003 to standardise OS field)

```

Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;
Query Match          99.8%; Score 1326; DB 2; Length 1329;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

Qy	1	TCAAAATCGAAGAGCTTTAGACTTGGCTTAACGCCCAAAAGAGGGGAAACCTGTTTATTTT	60
Db	1	TCAAAATCGAAGAGCTTTAGACTTGGCTTAACGCCCAAAAGAGGGGAAACCTGTTTATTTT	60
Qy	61	AGGGGAAGAATGCCTGTATGTAAATCTGGAATCAATTAATCTGAGAAAGTTAAAGA	120
Db	61	AGGGGAAGAATGCCTGTATGTAAATCTGGAATCAATTAATCTGAGAAAGTTAAAGA	120
Qy	121	AATTTGAGATCGAATATATAATGTAGACGAGAGCACTTCAAAAACATGCAACCTCGGGCT	180
Db	121	AATTTGAGATCGAATATATAATGTAGACGAGAGCACTTCAAAAACATGCAACCTCGGGCT	180
Qy	181	CCTCAGCCAAATGGATGCCCTGGACTCCCTCTTTAGGACCTCTAGCAGCTATAATATT	240
Db	181	CCTCAGCCAAATGGATGCCCTGGACTCCCTCTTTAGGACCTCTAGCAGCTATAATATT	240
Qy	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTAAAGTTTGTCTCTCCAGAAAT	300
Db	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTAAAGTTTGTCTCTCCAGAAAT	300
Qy	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCGAGTCATGACTAA	360
Db	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCGAGTCATGACTAA	360
Qy	361	AATCTACCGTGAACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAAATGACATTGA	420
Db	361	AATCTACCGTGAACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAAATGACATTGA	420
Qy	421	AGTCAACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTTACTACACTCCAAATTCAGTAGG	480
Db	421	AGTCAACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTTACTACACTCCAAATTCAGTAGG	480
Qy	481	AAGCAGTTTAGACGAGTTGTACGCCAACTCCCCAAACAGTACCTTGGGTTTTCTGTTGAGA	540
Db	481	AAGCAGTTTAGACGAGTTGTACGCCAACTCCCCAAACAGTACCTTGGGTTTTCTGTTGAGA	540
Qy	541	GGGTGACCTGACGACGAGGACTAGCTGGATTTTCTTAGCTGACTAGAAATCCCNAGCCT	600
Db	541	GGGTGACCTGACGACGAGGACTAGCTGGATTTTCTTAGCTGACTAGAAATCCCNAGCCT	600
Qy	601	ANCTGGGAAGGTGACCGGATCCCATCTTTAAACATGGGCTTGCGAACTTAGCTCACACCCG	660
Db	601	ANCTGGGAAGGTGACCGGATCCCATCTTTAAACATGGGCTTGCGAACTTAGCTCACACCCG	660
Qy	661	ACCAATCAGAGAGCTCACTAAAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC	720
Db	661	ACCAATCAGAGAGCTCACTAAAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC	720
Qy	721	AATCATCTATTCCCTGAGGACACGGGAAGGCAAGGATTTGGGATATAAACTCAGGCA	780
Db	721	AATCATCTATTCCCTGAGGACACGGGAAGGCAAGGATTTGGGATATAAACTCAGGCA	780
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATGTATGGGAGCTCTCTTTT	840
Db	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATGTATGGGAGCTCTCTTTT	840
Qy	841	CACCTCTATTCTCTATTAAATCATGCAACTGCACTCTTGGTCCGTGTTTTTTTATGG	900

[illegible]

RESULT 3

ADG14849	
ID	ADG14849 standard; CDNA; 1329 BP.
XX	
AC	ADG14849;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	MSRV associated CDNA #11.
XX	
KW	ss; pol gene; retrovirus; multiple sclerosis; rheumatoid arthritis.
XX	
OS	Multiple sclerosis associated retrovirus.
XX	
PN	US2003198647-A1.
XX	
PD	23-OCT-2003.
XX	
PF	03-APR-2002; 2002US-00114104.
XX	
PR	26-NOV-1996; 96US-00756429.
XX	
PR	26-NOV-1997; 97US-00979847.
XX	
XX	(INMR) BIO MERIEUX.
PA	
XX	
PI	Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
XX	
FI	Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
XX	
DR	WPI; 2004-032461/03.
XX	
XX	New isolated nucleic acid and their fragments having the pol gene of a
PT	retrovirus, useful for diagnosing, preventing and/or treating multiple
XX	
PT	sclerosis and/or rheumatoid arthritis.
XX	
XX	Disclosure; SEQ ID NO 108; 193pp; English.
PS	
XX	
CC	The invention relates to an isolated nucleic acid which comprises the pol
XX	
CC	gene of a retrovirus associated with multiple sclerosis or rheumatoid

CC arthritis. The methods and compositions of the present invention are
 CC useful for diagnosing, preventing and/or treating multiple sclerosis
 CC and/or rheumatoid arthritis. The present sequence is used in the
 CC exemplification of the invention.
 XX
 SQ Sequence 1329 BP; 346 A; 352 C; 288 G; 340 T; 0 U; 3 Other;
 Query Match 99.8%; Score 1326; DB 12; Length 1329;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACTGTATTATTTT 60
 DB 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGGGGAACTGTATTATTTT 60
 QY 61 AGGGGAAGAAGTCTTTAGTATGTTAAATCAATCTGGAATCAATCTAGAGAAGTTAAAGA 120
 DB 61 AGGGGAAGAAGTCTTTAGTATGTTAAATCAATCTGGAATCAATCTAGAGAAGTTAAAGA 120
 QY 121 AATTTGAGATCGAATATATGTTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGCGCT 180
 DB 121 AATTTGAGATCGAATATATGTTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGCGCT 180
 QY 181 CCTCAGCAATGGATGCGCTTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
 DB 181 CCTCAGCAATGGATGCGCTTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
 QY 241 TTTACTCTCTTTGGACCGCTGATCTTCAAACTTCTTGAAGTTTGTCTCTTCAGAAAT 300
 DB 241 TTTACTCTCTTTGGACCGCTGATCTTCAAACTTCTTGAAGTTTGTCTCTTCAGAAAT 300
 QY 301 TGAAGCTTAAAGCTTAAAGTAGTTCTTCAATGAAACCCAGAGTGCAGTCCATGACTAA 360
 DB 301 TGAAGCTTAAAGCTTAAAGTAGTTCTTCAATGAAACCCAGAGTGCAGTCCATGACTAA 360
 QY 361 AATCTACCGTGGACCGCTTGGACCGCTTGGACCTGCTAGACTATGCTGATGTTAATGACTGA 420
 DB 361 AATCTACCGTGGACCGCTTGGACCGCTTGGACCTGCTAGACTATGCTGATGTTAATGACTGA 420
 QY 421 AGTACCGCTCCGAGGAATCTCACTGCAACCCCTACTACTCACTCCATTCGATAGG 480
 DB 421 AGTACCGCTCCGAGGAATCTCACTGCAACCCCTACTACTCACTCCATTCGATAGG 480
 QY 481 AAGCAGTTTAGAGCAGTTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
 DB 481 AAGCAGTTTAGAGCAGTTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
 QY 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNNAAGCCT 600
 DB 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNNAAGCCT 600
 QY 601 ANCTGGGAAGGTACCGCATCGATCTTTAAACATGGGCTTGCACCTTAGCTCACCCCG 660
 DB 601 ANCTGGGAAGGTACCGCATCGATCTTTAAACATGGGCTTGCACCTTAGCTCACCCCG 660
 QY 661 ACCAATCAGAGCTCAGTAAATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 720
 DB 661 ACCAATCAGAGCTCAGTAAATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 720
 QY 721 AATCATCTATTGCTTACAGCAGCAGCGGAAGGACAAGGATTTGGGATATAAATCAGGCA 780
 DB 721 AATCATCTATTGCTTACAGCAGCAGCGGAAGGACAAGGATTTGGGATATAAATCAGGCA 780
 QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCATTTGATGGAGCTCTGTTTT 840
 DB 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCATTTGATGGAGCTCTGTTTT 840
 QY 841 CACTCTATTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTGTTTTTATGG 900
 DB 841 CACTCTATTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTGTTTTTATGG 900
 QY 901 CTCAGCTGAGCTTTTGTTCGCAATCCACACTGCTGTTTGGCCACCGTFCAGACCCGCT 960
 DB 901 CTCAGCTGAGCTTTTGTTCGCAATCCACACTGCTGTTTGGCCACCGTFCAGACCCGCT 960

DB 901 CTCAGCTGAGCTTTTGTTCGCAATCCACCACTGCTGTTTGGCCACCGTFCAGACCCGCT 960
 QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGGAGGT 1020
 DB 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGGAGGT 1020
 QY 1021 ACCCATTTGCCACTCCCGATCAGGCTTAAAGGCTTCCCAATTGTTCTGCAATGCTTAAGTCCC 1080
 DB 1021 ACCCATTTGCCACTCCCGATCAGGCTTAAAGGCTTCCCAATTGTTCTGCAATGCTTAAGTCCC 1080
 QY 1081 TGGGTTTTGTCTTAATAGAACTGAACACTGTGCTCACTGGGTTCCATTTCTTCCATGAC 1140
 DB 1081 TGGGTTTTGTCTTAATAGAACTGAACACTGTGCTCACTGGGTTCCATTTCTTCCATGAC 1140
 QY 1141 CCAGGCTTCTAATAGAGCTTAAACACTCACCAGATGCGCCCAAGATTCCATTTCTTGGTA 1200
 DB 1141 CCAGGCTTCTAATAGAGCTTAAACACTCACCAGATGCGCCCAAGATTCCATTTCTTGGTA 1200
 QY 1201 TCTGTGAGGCAAGAACCCAGGTCAGAGAANGTGGCTTGGCCACCAATTTGGGAAGTGG 1260
 DB 1201 TCTGTGAGGCAAGAACCCAGGTCAGAGAANGTGGCTTGGCCACCAATTTGGGAAGTGG 1260
 QY 1261 CCCACTGCCATTTTGGTAGCGGCCACCACTCTTGGAGCTGTGGAGCAAGGATCCC 1320
 DB 1261 CCCACTGCCATTTTGGTAGCGGCCACCACTCTTGGAGCTGTGGAGCAAGGATCCC 1320
 QY 1321 CCAGTAACA 1329
 DB 1321 CCAGTAACA 1329
 RESULT 4
 ADB84403
 ID ADB84403 standard; DNA; 1329 BP.
 XX
 AC ADB84403;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE MSRV-1 associated DNA sequence #21.
 XX
 KW ds; multiple sclerosis; rheumatoid arthritis; gag; pol;
 KW reverse transcriptase; ribonuclease H.
 XX
 OS Unidentified;
 XX
 FN US2003039664-A1.
 XX
 PD 27-FEB-2003.
 XX
 PP 26-NOV-1997; 97US-00979847.
 XX
 PR 26-NOV-1996; 96US-00756429.
 XX
 PA (PERR/) PERRON H.
 PA (BESE/) BESEME F.
 PA (BEDI/) BEDIN F.
 PA (PARA/) PARANHOS-BACCALA G.
 PA (KOMU/) KOMURIAN-PRADEL F.
 PA (JOLI/) JOLIVET-REYNAUD C.
 PA (MAND/) MANDRAND B.
 PA (GARS/) GARSON J A.
 PA (TUKU/) TUKU P W.
 XX
 PI Perron H, Beseme F, Bedin F, Paranhos-Baccala G;
 PI Komurian-Pradel F, Jolivet-Reynaud C, Mandrand B, Garson JA, Tuke PW;
 XX
 DR WPI; 2003-512253/48.
 XX
 PT New isolated or purified nucleic acid associated with multiple sclerosis
 PT and/or rheumatoid arthritis, useful for detecting a virus associated with
 PT multiple sclerosis or rheumatoid arthritis in a biological sample.
 XX

Claim 31; Page 80; 193pp; English.

PS The invention relates to an isolated or purified nucleic acid from a
 XX virus associated with multiple sclerosis and/or rheumatoid arthritis,
 CC multiple sclerosis-associated virus (MSRV)-1. The nucleic acids comprise
 CC pol, gag or reverse transcriptase genes (or their fragments) encoding the
 CC proteins or defined peptides (including immunodominant peptides,
 CC antigenic peptides or conserved motifs). Also included are a process for
 CC detecting a virus associated with multiple sclerosis or rheumatoid
 CC arthritis in a biological sample, a nucleic acid probe for the detection
 CC of a virus associated with multiple sclerosis or rheumatoid arthritis, a
 CC primer for the amplification by polymerization of a nucleic acid of a
 CC viral material associated with multiple sclerosis or rheumatoid
 CC arthritis, a polypeptide exhibiting an inhibitory activity on the
 CC proteolytic, reverse transcriptase or ribonuclease H activity from MSRV,
 CC and an antibody directed against the MSRV-1 virus obtained by
 CC immunologically reacting a human or animal body or cells with an
 CC immunogenic agent consisting of the antigenic polypeptide defined above.
 CC The nucleic acids are useful for detecting a biological sample, a virus
 CC associated with multiple sclerosis or rheumatoid arthritis, or for
 CC detecting in a biological sample, the presence of or exposure to a virus
 CC associated with multiple sclerosis or rheumatoid arthritis. The present
 CC sequence is a claimed MSRV-associated sequence whose identity cannot be
 CC accurately determined. Note: The SEQ ID numbers for the sequences as
 CC displayed in the main body of the patent do not match the SEQ ID numbers
 CC in the sequence listing. Consequently those sequences mentioned in the
 CC claims may not be the sequences the authors intended to claim.

XX SQ Sequence 1329 BP; 339 A; 346 C; 282 G; 332 T; 0 U; 30 Other;

Query Match 97.7%; Score 1299; DB 9; Length 1329;
 Best Local Similarity 98.0%; Pred. No. 0;
 Matches 1302; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTTATTTT 60
 DB 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTTATTTT 60
 QY 61 AGGGGAAGATCGTGTAGTATGTTAATCAATCTGGAATCACTTACGAGAAAGTAAAGA 120
 DB 61 AGGGGAAGATCGTGTAGTATGTTAATCAATCTGGAATCACTTACGAGAAAGTAAAGA 120
 QY 121 AATTTGAGATCGAATATAATGTAGACGAGGAGCACTTCAAAAACCTGCGCCCTGGGGCCT 180
 DB 121 AATTTGAGATCGAATATAATGTAGACGAGGAGCACTTCAAAAACCTGCGCCCTGGGGCCT 180
 QY 181 CCTCAGCCCAATGATGCCCTGGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATAT 240
 DB 181 CCTCAGCCCAATGATGCCCTGGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATAT 240
 QY 241 TTTTACTCCTCTTTGGACCCCTGTATCTCACTTCTTGTAGTTGTCTCTCCAGAAAT 300
 DB 241 TTTTACTCCTCTTTGGACCCCTGTATCTCACTTCTTGTAGTTGTCTCTCCAGAAAT 300
 QY 301 TGAAGCTGTAAAGCTCAAAATAGTCTTCAAAATGGAACCCAGATGCGATGCCATGACTAA 360
 DB 301 TGAAGCTGTAAAGCTCAAAATAGTCTTCAAAATGGAACCCAGATGCGATGCCATGACTAN 360
 QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTTG 420
 DB 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTTG 420
 QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGACACACCCCTACTACTCTCAATTCAGTAGG 480
 DB 421 AGTCACCCCTCCCGAGGAATCTCAACTGACACACCCCTACTACTCTCAATTCAGTAGN 480
 QY 481 AAGCAGTTAGCAGTGTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
 DB 481 AAGCAGTTAGCAGTGTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGN 540
 QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
 DB 541 GGGTGGACTGAGAGACAGGACTAGCTGGAATTTCTTAGGCTGACTAAGAAATCCCAAGCCN 600

QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCG 660
 DB ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCN 660
 QY 661 ACCAATCAGAGAGCTCTACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
 DB ACCAATCAGAGAGCTCTACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCN 720
 QY 721 AATCATCTATTGCTTCGAGAGCAGCGGGAAGCAAGGATTGGGATATATAAACTCAGGCA 780
 DB AATCATCTATTGCTTCGAGAGCAGCGGGAAGCAAGGATTGGGATATATAAACTCAGGCN 780
 QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGTATGGAGCTCTGTTT 840
 DB TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGTATGGAGCTCTGTTN 840
 QY 841 CACTCTATTGCTCTATTAAATCATGCACTGCACTCTCTGCTCCGTTTATG 900
 DB CACTCTATTGCTCTATTAAATCATGCACTGCACTCTCTGCTCCGTTTATGN 900
 QY 901 CTCAGCTGAGCTTTTGTTCGCCATCCACCTGCTGTTGGCACCGTCACAGACCCGCT 960
 DB CTCAGCTGAGCTTTTGTTCGCCATCCACCTGCTGTTGGCACCGTCACAGACCCGCN 960
 QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAGGT 1020
 DB GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGATCCAGCGAGNN 1020
 QY 1021 ACCCATGCTACCTCCGATCAGGCTTAAGGCTTGCCATTTGCTGCTGCTGCTTAAGTGCC 1080
 DB ACCCATGCTACCTCCGATCAGGCTTAAGGCTTGCCATTTGCTGCTGCTGCTTAAGTGNN 1080
 QY 1081 TGGGTTTGTCTTAATAGAACTGAACTGCTCACTGGGTTCCATGGTCTCTTCCATGAC 1140
 DB TGGGTTTGTCTTAATAGAACTGAACTGCTCACTGGGTTCCATGGTCTCTTCCATGNN 1140
 QY 1141 CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCCAAGATTCCTTCTTGGTA 1200
 DB CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCCAAGATTCCTTCTTGGNN 1200
 QY 1201 TCTGTGAGCCCAAGAACCCCGAGTCAGAGAANGTAGGCTTGCCACCATTTGGGAAGTGG 1260
 DB TCTGTGAGCCCAAGAACCCCGAGTCAGAGAANGTAGGCTTGCCACCATTTGGGAAGTNN 1260
 QY 1261 CCCACTGCCATTTGTAGCGGCCACACCATCTTTGGGAGCTGTGGGAGCAAGGATCCC 1320
 DB CCCACTGCCATTTGTAGCGGCCACACCATCTTTGGGAGCTGTGGGAGCAAGGATCNN 1320
 QY 1321 CCAGTAACA 1329
 DB CCAGTAACA 1329

RESULT 5
 ABL61744

ID ABL61744 standard; DNA; 56093 BP.

XX AC ABL61744;

XX XX

DE 15-MAY-2002 (first entry)

XX Colon adenocarcinoma related gene sequence SEQ ID NO:81.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 XX gene; ds.

OS Homo sapiens.

XX WO200194629-A2.

XX

PD 13-DEC-2001.
XX 30-MAY-2001; 2001WO-US010838.
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-02331133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 22-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 28-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.

(AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 81; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,

CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 56093 BP; 16164 A; 12346 C; 10702 G; 16881 T; 0 U; 0 Other;

Query Match 77.9%; Score 1035.2; DB 6; Length 56093;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;

Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTACCGCCAAAGAGGGGGAACCTGTTTATTTT 60
Db 37018 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGGAACCTGTTTATTTT 37077

Qy 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGATCATTTACTCAGAAAGTTAAAGA 120
Db 37078 AGGGGAAGATGCTGTTATTTATTTAATCAATCGGAATCGTCACTGAGAAAGTTAAAGA 37137

Qy 121 AATTTGAGATCGAATATAATGTAGAGCAGAGGAGCTTCAAAACACTGCACCTGGGCGCT 180
Db 37138 AATTCGAGATCGAATACAACTGAGAGCAGAGGAGCTTCGAAACACTGGACCTGGGCGCT 37197

Qy 181 CCTCAGCCAAATGGATGCCCTGGACTCTCCCTTTCTTAGGACCTTAGCAGCTATAATTT 240
Db 37198 CCTCAGCCAAATGGATGCCCTGGATTTCTCCCTTTCTTAGGACCTTAGCAGCTATAATTT 37257

Qy 241 TTTACTCTCTTTGGACCCCTGTATCTTCAACTCTCTTTAAAGTTTGTCTCTCCAGAAAT 300
Db 37258 GCTACTCTCTTTGGACCCCTGTATCTTAACTCTCTTTAACTTTGTCTCTCCAGNAT 37317

Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db 37318 CGAAGCTGTAAGCTA-----CAAAATGGAGCCCAAGATGCAGTCCAAAGACTAA 37365

Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTTGA 420
Db 37366 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCCGATCTGATGTTAATGACATCAA 37425

Qy 421 AGTCACCCCTCCGAGGAAATCTCACTGCAACACCCCTACTACATCTCAATTCAGTAGG 480
Db 37426 AGGCACCCCTCTCAGGAAATCTCAGCTGCAACACCTCTACTAGCCCAATTCAGCAGG 37485

Qy 481 AAGCAGTTAGACGAGTTGTCTAGCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
Db 37486 AAGCAGTTAGACGCTGTCTGGCCAACTCTCCCAACAGCACTAGGTTTTCTGTTGAGA 37545

Qy 541 GGGTGGACTGAGACAGGACTAGTGGATTTCTTAGCTGACTAAGAAATCCNAAGCCT 600
Db 37546 TGGGGGACTGAGACAGGACTAGTGGATTTCTTAGCTGACTAAGAAATCCCTAAGCCT 37605

Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 37606 AGCTGGGAGGTGACCATCCACTTTAAACAGGGGCTTGCAACTTAGCTCACACCTG 37665

Qy 661 ACCAATCAGAGAGTCTCAATAATGCTAATCAGGCAAAAACAGGAGTTAAGCAATAGCC 720
Db 37666 ACCAATCAGAGAGTCTCAATAATGCTAATAGGCAAAACAGGAGGTTAAGCAATAGCC 37725

Qy 721 AATCATCTATTGCTGAGACAGCGGAGGACAGGATTTGGGATATAAATCAGGCA 780
Db 37726 AATCATCTATTGCTGAGACAGCGGAGGAGGACATGATCGGGATATAAATCCCAATC 37785

Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
Db 37786 TTCAAGCAGCAACAGGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTT 37845

Qy 841 CACTCTATTTCATCTATTAATCATGCAACTGCACTCTTCTGGTCCCTGTTTTATGG 900
Db 37846 CATGCTATTTCACTCTATTAATCTTGCAACTGCACTCTTCTGGTCCATGTTTCTACGG 37905

Qy 901 CTCAGCTGAGCTTTGTTCCGCACTCCACCACTGCTGTTTGGCACCGTCACAGACCGCT 960
Db

Db 37906 CTTGAGCTGAGCTTTGCTGCGCATCCACCACTGCTGTTTTCGCGCCACGCGAGCCCGCC 37965
 QY 961 GTGACTTCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCGAGGT 1020
 Db 37966 GTGACTCCCATCCCTCTGGATCATGACAGGGTGTGCTGCTGATCCAGCGAGGC 38025
 QY 1021 ACCATTGGCACTCCCGATCAGCTAAAGCTTGGCAATTTCTTCATGGCTTAAGTGCC 1080
 Db 38026 ACCATTGGCGCTCCCAATCGGGCTAAAGCTTGGCAATTTCTTCATGGCTTAAGTGCC 38085
 QY 1081 TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTTCCATGAC 1140
 Db 38086 TGGGTTTCATCTTAATTGAGCTGAACACTAGTCACTGGGTTCCATGGTTCTCTTCTGTGAC 38145
 QY 1141 CACGGCTTCTAATAGACTATTAACACTCAGCGCATGGCCCAAGATTCATTCTCTTGGTA 1200
 Db 38146 CCACAGCTTCTAATAGACTATTAACACTCAGCGCATGGCCCAAGATTCATTCTCTT-GAA 38204
 QY 1201 TCTGTGAGCCCAAGAACCCCGCTCAGAGAANGTGAGGCTTGGCACCATTTGGGAAG 1257
 Db 38205 TCCATAGGCCCAAGAACCCCGCTCAGAGAANGTGAGGCTTGGCACCATTTGGGAG 38261

RESULT 6

ADFS9718
ID ADFS9718 standard; cDNA; 9502 BP.

XX AC
XX AC
XX AC

DT 12-FEB-2004 (first entry)

XX Human contig polynucleotide sequence SEQ ID NO:2085.

KW biological activity; genetic engineering; hybridisation probe; oligomer;
 KW primer; chromosome mapping; gene mapping; recombinant protein production;
 KW human; gene; ss.

XX Homo sapiens.

XX WO2003080795-A2.

XX 02-OCT-2003.

XX 09-AUG-2002; 2002WO-US025485.

XX 09-AUG-2001; 2001US-0311261P.

XX (HYSB-) HYSBQ INC.

XX Tang YT, Yang Y, Wang Z, Weng G, Ma Y;

XX WPI; 2003-876918/81.

XX P-PSDB; ADF60170.

XX New polynucleotides, useful as hybridization probes, oligomers or
 PT primers, for chromosome or gene mapping, for the recombinant production
 PT of proteins, and for generating antisense DNA or RNA.

XX Example 2; SEQ ID NO 2085; 571bp; English.

XX The present invention describes isolated polynucleotide sequences (I),
 CC which encode polypeptides (II) with biological activity. Also described:
 CC (1) a vector comprising (I); (2) an expression vector comprising (I); (3)
 CC a host cell genetically engineered to comprise (I) which is operatively
 CC associated with a regulatory sequence that modulates expression of (I) in
 CC the host cell; (4) a polypeptide (II) encoded by (I); (5) a composition
 CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed
 CC against the polypeptide of (4); (7) detecting (I) or the polypeptide of
 CC (4) in a sample; (8) identifying a compound that binds to the polypeptide
 CC of (4); (9) producing the polypeptide of (4); and (10) a collection of
 CC polynucleotides comprising at least one of the polynucleotide sequences
 CC (I). The polynucleotides (I) can be used as hybridisation probes,
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant

CC production of proteins, and for generating antisense DNA or RNA. The
 CC present sequence represents a human contig polynucleotide sequence, which
 CC is used in an example from the present invention.

XX Sequence 9502 BP; 2813 A; 2433 C; 1992 G; 2263 T; 0 U; 1 Other;

SQ Query Match 77.5%; Score 1029.4; DB 10; Length 9502;
 Best Local Similarity 87.9%; Pred. No. 0;
 Matches 1178; Conservative 0; Mismatches 129; Indels 33; Gaps 4;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTTAACCCGCAAAAGAGGGGAACTGTTATTTT 60
 Db 2575 TCAAAATCGAAGAGCTTTAGACTTCTTAACCCGCTGAAAGAGGGGAACTGTTATTTT 2634
 QY 61 AGGGGAAGAATGCTGTAGTAAATCAATCTGAATCATTTACTGAGAAAGTTAAAGA 120
 Db 2635 AGGGGAAGAATGCTGTAGTAAATCAATCTGAATCATTTACTGAGAAAGTTAAAGA 2694
 QY 121 AATTTGAGATCGAATATATAATGTAGAGCAGAGGACCTTCAAAAACACTGCAACCTGGGGCCT 180
 Db 2695 AATTCGAGATCGAATATACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCTGGGGCCT 2754
 QY 181 CCTCAGCAATGGATGCCCTCGACTCTCCCTTTTAGGACCTCTAGCAGCTATAATAT 240
 Db 2755 CCTCAGCAATGGATGCCCTCGACTCTCCCTTTTAGGACCTCTAGCAGCTATAATAT 2814
 QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTGTCTTCCAGAAT 300
 Db 2815 GCTACTCTCTTTGGACCTGTATCTTAAACCTCTTGTAAAGTTGTCTTCCAGAAT 2874
 QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCCGAGATGCACTCAATGACTAA 360
 Db 2875 CGAAGCTGTAAACTA-----CAAAATGGAACCCCGAGATGCACTCAATGACTAA 2922
 QY 361 AATCTACCGTGAACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAAAGACTGA 420
 Db 2923 GATCTACCGCAGACCCCTGGACCGGCTGTAGACTGATCTGATGTTAAAGACTGA 2982
 QY 421 AGTCAACCTCCCGAGGAATCTCAACTGCAACCCCTACTACACTCAATTCAGTAGG 480
 Db 2983 AGGCACCCCTCTCTGAGGAATCTCAAGCTGCAACCTCTACTAGCCCCCAATTCAGCAGG 3042
 QY 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA 540
 Db 3043 AAGCAGTTAGAGCGGTGCTGCGCCCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA 3102
 QY 541 GGGTGAAGTGAAGACAGACTAGCTGGATTTCTAGGCTGACTAAGATCCCNAGCCT 600
 Db 3103 TGGGGGACTGAGAGCAGGACTAGCTGGATTTCTAGGCTGACTAAGATCCCNAGCCT 3162
 QY 601 ANCTGGGAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660
 Db 3163 AGCTGGGAAGGTAAACCATCCACTTTAAACACGGGGCTTGCACCTTAGCTCACACCCG 3222
 QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGGAGGTAAA 711
 Db 3223 ACCAATCAGGTAGTAAAGAGAGCTCACTAAATGCTTAATAGGCAAAAACAGGAGGTAAA 3282
 QY 712 GCAATAGCCAAATCATCTATTGCTGAGACACAGCGGGAAGGACAGGATTTGGATATAA 771
 Db 3283 GAAATAGCCAAATCATCTATTGCTGAGACACAGCGGGAAGGACAGGATTTAGATATAA 3342
 QY 772 ACTCAGGCAATTCAGCCAGCAGCAACCCCTTTGGGTCCCTCCCTCCCTGATGATGGAG 831
 Db 3343 ACCCAGGCAATTCAGCCAGCAGCAGGCTACCCCTTTTGGGTCCCTCCCTGATGATGGAG 3402
 QY 832 CTCTGTTTTCACTCTATTTCACTCTATTAAATCATGCAACTGCA--CTCTTCTGGTCCGT 889
 Db 3403 CTCTGT-----TTCCACTCTATTAAATCTTGCNACTGCACCTCTCTCTGGTCCGT 3452
 QY 890 GTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATTCACCACTGCTGTTTGGCCACCGTC 949
 Db 3453 GTTTGTTACGGATTGAGCTGAGCTTTTGTGCTTATTGTTCCACCACTTGTCTTTGGCGCTGT 3512

QY 950 ACAGACCGCTGCTGACTTCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGA 1009
 DB 3513 GCAGACCTGCCACTGACTTCCATCCCTCTGGATCCGAGGGTGTCCGCTGTGCTCTGA 3572
 QY 1010 TCCAGCGAGGTACCATTTGCCACTCCGATCAGGCTAAAGGCTTGCCATTGTTCTCTGCAT 1069
 DB 3573 TCCAGCGAGGCGCCATTGCCCTCCGATCGGGCTAAAGGCTTGCCATTGTTCTCTGCAC 3632
 QY 1070 GGCTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCTACTGGTTCCATGGTTC 1129
 DB 3633 GGCTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCTACTGGTTCATGGTTC 3692
 QY 1130 TCTTCATGACCCAGCGGCTTCTAATAGAGCTATTAACACTCACCGCATGGCCCAAGATTCC 1189
 DB 3693 TCTTCATGACCCAGCGGCTTCTAATAGAGCTATTAACACTCACCGCATGGCCCAAGATTCC 3752
 QY 1190 ATTCTCTGGTATCTGTAGGCGCAAGACCCAGGTTCAGAGAAAGTGGAGCTTGGCCACCAT 1249
 DB 3753 ATTCTCTGGTATCTGTAGGCGCAAGACCCAGGTTCAGAGAAAGTGGAGCTTGGCCACCAT 3812
 QY 1250 TTGGGAAGTGGCCACTGCCATTTTGGTATGGCGGCCACCACTCTTGGGAGCTGTGGGA 1309
 DB 3813 CTGGGAAGTGGCCACTGCCATTTTGGTATGGCGGCCACCACTCTTGGGAGCTGTGGGA 3872
 QY 1310 GCAAGGATCCCGCAGTAACA 1329
 DB 3873 GCAAGGATCCCGCAGTAACA 3892

RESULT 7

ABN97929 standard; DNA; 10499 BP.

XX AC ABN97929;

XX 01-AUG-2002 (first entry)

XX Human retroviral sequence HERV-7q.

XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;

XX multiple sclerosis; ds.

XX Human endogenous retrovirus.

XX WO967395-A1.

XX 29-DEC-1999.

XX 23-JUN-1999; 99WO-FR001513.

XX 23-JUN-1998; 98FR-00007920.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Alliel PM, Perin J, Rieger F;

XX WPI; 2000-160587/14.

XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used

XX for diagnosis, treatment and prevention of autoimmune and neurological

XX diseases.

XX Claim 3; Fig 1; 225pp; French.

XX The present invention relates to new nucleic acid sequences of human

XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.

XX Regulatory elements associated with HERV-7q may alter expression of other

XX genes (even remote genes) on the same chromosome, inducing immunological

XX and/or neurological changes (which may be pathological or protective/

XX curative). HERV-7q peptides can be used to improve efficiency of the

XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding

XX sequences can be used in immunogenic or vaccinating compositions, for

CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention

XX SQ Sequence 10499 BP; 3048 A; 2576 C; 2280 G; 2495 T; 0 U; 0 Other;

Query Match 77.0%; Score 1023.2; DB 3; Length 10499;
 Best Local Similarity 90.1%; Pred. No. 0;
 Matches 1132; Conservative 0; Mismatches 111; Indels 14; Gaps 3;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTATTATTTT 60
 DB 9018 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAAGAGGGGAACTGTATTATTTT 9077
 QY 61 AGGGGAAGATGCTGTGTATGTTAATCAATCTGGAATCATTTACTGAGAAGTTAAGA 120
 DB 9078 AGGGGAAGATGCTGTGTATGTTAATCAATCGGAATCGTCACTGAGAAGTTAAGA 9137
 QY 121 AATTTGAGATCGAATAATATGATAGCAGAGGACCTTCAAAACACACTGCACCTGGGCGCT 180
 DB 9138 AATTCGAGATCGAATAACAGTAGCAGAGGAGCTTCGAAACACTGGAACCTGGGCGCT 9197
 QY 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATT 240
 DB 9198 CCTCAGCAATGGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATT 9257
 QY 241 TTTACTCCTCTTTGGACCTGTATCTCAACTTCTTGTAAAGTTCTCTCTCCAGAAAT 300
 DB 9258 GCTACTCCTCTTTGGACCTGTATCTTTAACTCTCTTTAACTTTCTCTCTCCAGAAAT 9317
 QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAGTCCATGACTAA 360
 DB 9318 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCAGTCCCAAGACTAA 9365
 QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTGTAGACTATGCTCTGATTTAATGACATTGA 420
 DB 9366 GATCTACCGGACAGCCCTGGACCGGCTGTGTAGCCACGATCTGATGTTAATGACATCAA 9425
 QY 421 AGTCACCCCTCCGAGGAATCTCAATGTCACACACCCCTACTACATCCCAATCCAGTAGG 480
 DB 9426 AGGCACCCCTCCTGAGGAATCTCAGCTGCAACAACCTCTACTAGCCCAATCCAGTAGG 9485
 QY 481 AAGCAGTTAGAGCAGTTGTGAGCCAACTCTCCCAACAGTACTTGGGTTTTCTGTGTAGA 540
 DB 9486 AAGCAGTTAGAGC--GGTCTCGGCCAACTCTCCCAACAGCAGCTTAGGTTTTCTGTGTAGA 9544
 QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGGCT 600
 DB 9545 TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGGCT 9604
 QY 601 ANCTGGGAAGTGCACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCGG 660
 DB 9605 AGCTGGGAAGTGCACCGCATCCATCCCTTTAAACACGGGGCTTGCAACTTAGCTCACACCG 9664
 QY 661 ACCAATCAGAGCTCAGTAAATGCTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC 720
 DB 9665 ACCAATCAGAGCTCAGTAAATGCTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC 9724
 QY 721 AATCATCTATTGCTGAGAGACAGCGGAAGGACAGGATTTGGGATATAAATCACTCAGGCA 780
 DB 9725 AATCATCTATTGCTGAGAGACAGCGGAAGGCAATGATCGGGATATAAATCACTCAGGCA 9784
 QY 781 TTCAAGCAGCAACAGCAACCCCTTTGGGCTCCCTCCATTTGATGGAGCTCTGTATT 840
 DB 9785 TTCAAGCAGCAACAGCAACCCCTTTGGGCTCCCTCCATTTGATGGAGCTCTGTATT 9844
 QY 841 CACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTATGG 900
 DB 9845 CATGCTATTTCATCTATTAAATCTGCAACTGCACTCTTCTGGTCCATGTTTCTTACGG 9904
 QY 901 CTCAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTGTTGCCACCGTCACAGACCGCT 960

DT 04-DEC-2001 (first entry)

XX Human diagnostic and therapeutic polynucleotide (DITHP) #17.

DE Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;

KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;

KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;

KW respiratory disorder; ss.

XX Homo sapiens.

OS WO200162927-A2.

PN 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US006059.

XX 24-FEB-2000; 2000US-0184693P.

XX 24-FEB-2000; 2000US-0184697P.

XX 24-FEB-2000; 2000US-0184698P.

XX 24-FEB-2000; 2000US-0184768P.

XX 24-FEB-2000; 2000US-0184769P.

XX 24-FEB-2000; 2000US-0184770P.

XX 24-FEB-2000; 2000US-0184771P.

XX 24-FEB-2000; 2000US-0184772P.

XX 24-FEB-2000; 2000US-0184773P.

XX 24-FEB-2000; 2000US-0184774P.

XX 24-FEB-2000; 2000US-0184776P.

XX 24-FEB-2000; 2000US-0184777P.

XX 24-FEB-2000; 2000US-0184797P.

XX 24-FEB-2000; 2000US-0184813P.

XX 24-FEB-2000; 2000US-0184837P.

XX 24-FEB-2000; 2000US-0184841P.

XX 24-FEB-2000; 2000US-0185213P.

XX 24-FEB-2000; 2000US-0185216P.

XX 12-MAY-2000; 2000US-0203785P.

XX 15-MAY-2000; 2000US-0204226P.

XX 16-MAY-2000; 2000US-0204525P.

XX 16-MAY-2000; 2000US-0204821P.

XX 16-MAY-2000; 2000US-0204908P.

XX 16-MAY-2000; 2000US-0205232P.

XX 17-MAY-2000; 2000US-0204815P.

XX 17-MAY-2000; 2000US-0204863P.

XX 17-MAY-2000; 2000US-0205221P.

XX 17-MAY-2000; 2000US-0205285P.

XX 17-MAY-2000; 2000US-0205286P.

XX 17-MAY-2000; 2000US-0205287P.

XX 17-MAY-2000; 2000US-0205323P.

XX 17-MAY-2000; 2000US-0205324P.

XX (INCY-) INCYTE GENOMICS INC.

XX Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;

PI Chen A, D'ea SA, Amesby S, Dahl CR, Dam TC, Daniels SE, Dufour GE;

PI Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL, Liu TF;

PI Roseberry AM, Rosen BR, Russo FD, Stockdreher TK, Daffo A;

PI Wright RJ, Yap PE, Yu JV, Bradley DL, Bratcher SR, Chen W;

PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;

XX WPI; 2001-502867/55.

DR P-PSDB; AAU19431.

XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.

PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.

XX Claim 1; Page 304; 522pp; English.

XX The invention relates to polynucleotides (I) encoding diagnostic and

CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and

CC proteins involved in growth and development and receptors. (I) and (II)

CC may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate DITHP expression. For example, (I) and (II)

CC may be used to treat disorders associated with decreased polypeptide

CC expression by rectifying mutations or deletions in a patient's genome,

CC that affect the activity of the DITHPs, by expressing inactive proteins

CC or supplementing the patient's own production of them. (I) and (II) may

CC be used to treat diseases, for example, cell proliferative disorder,

CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,

CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,

CC (I) may be used to produce the DITHPs, by inserting the nucleic acids

CC into a host cell and culturing the cell to express the protein. (I) and

CC its complementary sequences may also be used as DNA probes in diagnostic

CC assays to detect and quantitate the presence of similar nucleic acids in

CC samples, and therefore which patients may be in need of restorative

CC therapy. (II) may also be used as antigens in the production of DITHP

CC antibodies against DITHPs and in assays to identify modulators of DITHP

CC expression and activity. The anti-DITHP antibodies and antagonists may

CC also be used to down regulate expression and activity. The anti-DITHP

CC antibodies may also be used as diagnostic agents for detecting the

CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay

CC (ELISA)). AAS30986-AAS31196 represent human diagnostic and therapeutic

CC (DITHP) polynucleotides of the invention

XX

SQ Sequence 1393 BP; 375 A; 352 C; 304 G; 360 T; 0 U; 2 Other;

Query Match 58.6%; Score 778.8; DB 4; Length 1393;

Best Local Similarity 87.2%; Pred. No. 1.2e-244;

Matches 904; Conservative 0; Mismatches 112; Indels 21; Gaps 4;

QY 302 GAAGCTGTAAGCTACAAATAGTTCTTCAAATGGAAACCCAGATGCGATGACTATAA 361

DB 1 GAAGCTGTAAGCTACAAATAGTTCTTCAAATGGAAACCCAGATGCGATGACTATAA 60

QY 362 ATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACTGAA 421

DB 61 ATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACTGAA 120

QY 422 GTACCCCTCCGAGGAAATCTCAATGCGACACCCCTACTACCTCCATTCAGTAGGA 481

DB 121 AGCACTCC-CCCAAGGAAATTTCACTGCAACACCCCTACTACACCCAAATCAGCAGGA 179

QY 482 AGCAGTTAGAGCAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTTCTCTGTGAGAG 541

DB 180 AGCAGTTAGAGCAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTTCTCTGTGAGAG 239

QY 542 GGTGGAAGTGACCGGCTAGCTGGATTTCTAGGCTGACTAAGAAATCCCAAGGCTA 601

DB 240 CGGGGACTGAGACAGGAGTACTGCTGATTTCTAGGCTGACTAAGAAATCCCAAGGCTA 299

QY 602 NCTGGGAAGTGACCGGCTAGCTGGATTTCTAGGCTGACTAAGAAATCCCAAGGCTA 661

DB 300 GCTGGGAAGTGACCGGCTAGCTGGATTTCTAGGCTGACTAAGAAATCCCAAGGCTA 359

QY 662 CCAATC-----AGAGAGCTCACTAAATGCTAAATCAGGCAAAACAGGAGGTAAAG 712

DB 360 CCAATCAGGTAGTAAAGAGAGCTCACTAAATGCTAAATGCTAAATGCTAAATGCTAAAG 419

QY 713 CAATAGCCCAATCATCTATTGCTGAGACGACGCGGGAAGGACAAGGATTTGGGATATAA 772

DB 420 AATAGCCCAATCATCTATTGCTGAGACGACGCGGGAAGGACAAGGATTTGGGATATAA 479

QY 773 CTCAGGCATTCAAGCAGCAACACCCCTTTGGGTCCTCCCTCCCACTGTTATGGGAGC 832

DB 480 CCAAGCATTTCGAGCAGCAACACCCCTTTGGGTCCTCCCTCCCACTGTTATGGGAGC 539

QY 833 TCTGTTTTCCTACTATTTCCTACTTAAATCATGCAACTGCACTCTTCTGGTCCGTGTT 892

DB 540 TCTGT-----TTTCACTCTATTAAATCTTGAACGCTGCACTCTTCTGGTCCGTGTT 589

QY 893 TTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACGTCACA 952

DB 590 TGTTCGGTTTGGCTGAGCTGAGCTTTTCGCTGCGCGTCCACCACTGCTGTTTGGCGCATCGGA 649

QY 953 GACCGCTGCTGATCTCCATCCCTTTGGATTCAGCAGAGGTCCACCTGCTGCTGATCC 1012

DB 650 GACCTGCGCTGACTTTCATCCCTCCCGATCTGCGAGGGTGTTCAT-TGTGCTCTGATCC 708

```
QY 1013 AGCAGGTACCATTTGCCACTCCCGATCAGCTAAAGCCTTGCCATTGTTCTCGATGGC 1012
DB |||||
QY 1073 TAAGTGCTGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCT 1132
DB |||||
QY 1133 TCCATGACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCAT 1192
DB |||||
QY 1193 CTTGGTATCTGTAGGCGCAAGAACCCAGCTCAGAGAAAGTGAGGCTTGCCACCATTTG 1252
DB |||||
QY 1253 GGAAGTGCCCACTGCCATTTTGGTAGCGGCCACCACTCTTGGGAGCTGTGGAGCA 1312
DB |||||
QY 1313 AGGATCCCCCAGTAACA 1329
DB |||||
QY 1009 AGGACCCCTGTGAACA 1025
DB |||||

RESULT 11
AAA63826
ID AAA63826 standard; DNA; 2030 BP.
XX
AC AAA63826;
XX
DT 06-AUG-2003 (revised)
DT 04-DEC-2000 (first entry)
XX
DE Nucleotide sequence of the MSRV-1 3' env and LTR regions.
XX
KW MSRV-1; pol region; long terminal repeat; LTR; RUS region; retrovirus;
KW ss.
XX
OS Multiple sclerosis associated retrovirus.
XX
PH Key Location/Qualifiers
FT CDS 1..1629
FT /tag= a
FT /note= "Contains one termination codon"
FT sig_peptide 1..81
FT /tag= b
FT CAAT_signal 1800..1807
FT /tag= c
FT CAAT_signal 1858..1864
FT /tag= d
FT TATA_signal 1906..1911
FT /tag= e
FT polyA_signal 1996..2002
FT /tag= f
XX
PN WO200047745-A1.
XX
PD 17-AUG-2000.
XX
PF 15-FEB-2000; 2000WO-IB0000159.
XX
PR 15-FEB-1999; 99EP-00420041.
XX
PA (INNR ) BIO MERIEUX.
XX
PI Paranhos-Baccala G, Perron H, Komurian-Pradel F;
XX
DR WPI; 2000-506097/45.
DR P-PSDB; AAB08195.
XX
PT Nucleotide fragment of LTR-RUS region from Multiple Sclerosis retrovirus
```

```
PT (MSRV) used to detect the presence of MSRV-1 retrovirus in a biological
PT sample.
XX
XX Disclosure; Fig 2; 23pp; English.
XX
CC The present sequence represents the nucleotide sequence corresponding to
CC the 3' env region and long terminal repeat sequences from clone CL6 of
CC Multiple Sclerosis retrovirus (MSRV-1). The specification describes a
CC long terminal repeat (LTR)-RUS region which encodes the expression of a
CC MSRV-1 protein. This is unusual for LTRs, in particular in the RUS
CC region. The sequence includes CAAT and TATA signals which are present in
CC the U3 and R regions and are not directed towards the CPS indicated in
CC the features table. Probes and antibodies to the MSRV-1 retrovirus
CC protein and encoding polynucleotide sequences are used to detect the
CC presence of MSRV-1 retrovirus in a biological sample. (Updated on 06-AUG-
CC 2003 to correct OS field.)
XX
SQ Sequence 2030 BP; 574 A; 559 C; 387 G; 510 T; 0 U; 0 Other;
Query Match 58.3%; Score 775.4; DB 3; Length 2030;
Best Local Similarity 92.8%; Pred. No. 2.1e-243;
Matches 812; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 1 TCAAAATCGAAGAGCTTTAGACTTTGTAACCGCCAAAGAGGGGAACTGTTATTTT 60
DB |||||
QY 61 AGGGGAAGATGCTGTAGTATGTTAATCAATCTGGATCACTTCTGGAAGTAAAGA 120
DB |||||
QY 1200 AGGAGAAGACGCTGTTATTATGTTAATCAATCAGAAATGTCATCGAAGAAGTTAAAGA 1259
DB |||||
QY 121 AATTGAGATCGAATAATAATGAGCAGAGAGACCTTCAAAACACATGACACCTGGGGCT 180
DB |||||
QY 1260 AATTCGAGATCGAATACAATGAGCAGAGAGCTTCAAAACACCGAAGCTGGGGCT 1319
DB |||||
QY 181 CCTCAGCAATGGATGCCCTGGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATT 240
DB |||||
QY 1320 CCTCAGCAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATT 1379
DB |||||
QY 241 TTACTCTCTCTTGGACCTGTATCTCACTTCACTCTCTTAAAGTTTCTCTCCAGAA 300
DB |||||
QY 1380 GTTACTCTCTCTTGGACCTGTATCTTAAACCTCTCTTAAAGTTTCTCTCCAGAA 1439
DB |||||
QY 301 TGAAGCTGTAAGACTACAATAGTTCTTCAAATGGAACCCAGATGCACTGACACTAA 360
DB |||||
QY 1440 TGAAGCTGTAAGACTACAAGCTACAGATGGTCTTACAATGGAACCCAGATGCACTGAC 1499
DB |||||
QY 361 AATCTACCGTGGAACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
DB |||||
QY 1500 GATCCACGCTGGACCCCTGGACCGGCTGTAGCCCATGCTCCGATGTTAATGACATTGA 1559
DB |||||
QY 421 AGTCACCCCTCCCGAGGAATCTCACTGCAACAACCCCTACTACACTCCAATTCAGTAGG 480
DB |||||
QY 1560 AGGCACCCCTCCCGAGGAATCTCACTGCAACAACCCCTACTATGCCCAATTCAGCGGG 1619
DB |||||
QY 481 AAGCAGTTAGAGCAGTTGTGAGCCAAACCTCCCAACAGTACTTGGGTTTTCTGTGAGA 540
DB |||||
QY 1620 AAGCAGTTAGAGCGGTCAATCAGCCAACTCCCAACAGCACTTGGTTTTCTGTGAGA 1679
DB |||||
QY 541 GGGTGGAAGTGAAGGAGTCTGATGTTTCTAGGCTGACTAAGAAATCCCNAGGCT 600
DB |||||
QY 1680 GGGGGGACTGAGAGACAGGACTAGCTGGAATTTCTTAGGCCAACGAAGAAATCCCTAAGGCT 1739
DB |||||
QY 601 ANCTGGGAAGGTGACCGCATCTTTTAAACATGGGGCTTGCACCTTAGCTCACACCCG 660
DB |||||
QY 1740 AGCTGGGAAGGTGACTGATCCACTCTAAACATGGGGCTTGCACCTTAGCTCACACCCG 1799
DB |||||
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC 720
DB |||||
QY 1800 ACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAAATAGGAGGTAAGCAATAGCC 1859
DB |||||
QY 721 AATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATTGGGATATAAAGTCAAGCA 780
DB |||||
```

Db 1860 AATCATCTATTGCTGAGAGACAGCGGAGGAGCAAGAGTCGGGATATAAAACCCAGGCA 1919
Qy 781 TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCTCCCAATGTTATGGGAGCTCTGTTTT 840
Db 1920 TTGAGCCGCGCAACGCGCAACCCCTTTGGTCCCTCCCTCCCTTTGTTATGGGCGCTCTGTTTT 1979
Qy 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGCA 875
Db 1980 CACTCTATTTCACCTCTATTAAATCTTGCAACTGAA 2014
RESULT 12
ACN88786
ID ACN88786 standard; DNA; 7974 BP.
XX ACN88786;
XX
XX
XX 02-DEC-2004 (first entry)
XX
XX Breast cancer related marker, seq id 9936.
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX Homo sapiens.
XX OS
XX US2003099974-A1.
XX
XX 29-MAY-2003.
XX
XX 18-JUL-2002; 2002US-00198846.
XX
XX 18-JUL-2001; 2001US-0306220P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2003-787014/74.
XX
XX Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.
XX
XX Disclosure; SEQ ID NO 9936; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with
XX breast cancer which is encoded by a nucleic acid molecule comprising a
XX nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX the polypeptide of the invention. The activity of the polypeptide of the
XX invention may be described as cytostatic. The antibody is useful for
XX detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX invention are useful in the detection of breast tumours. (I) is useful as
XX a marker for breast cancer and in breast cancer therapy. Sequences given
XX in records ACN78851-ACN92934 represent nucleic acid markers associated
XX with breast cancer. Note: The sequence listing does not form part of the
XX specification but may be obtained in electronic format from the USPTO web
XX site at seqdata.uspto.gov/sequence.html?DocID=2003099974
XX
XX Sequence 7974 BP; 2034 A; 2017 C; 2123 G; 1800 T; 0 U; 0 Other;
Query Match 58.38; Score 774.2; DB 11; Length 7974;
Best Local Similarity 85.8; Pred. No. 1.2e-242;
Matches 937; Conservative 0; Mismatches 126; Indels 29; Gaps 6;
Qy 257 CCCTGTATCTTCAACTTCCTTGTAAAGTTGTCTCTTCCAGAAATGAAGCTGTAAAGCTA 316
Db 6409 CCCTGTATCTTAACTCTCTTGTAAAGTTGTCTCTTCCAGAAATGAAGCTGTAAAGCTA 6468
Qy 317 CAAATAGTCTTCAAAATGGAAACCCAGATGAGTCCATGAC-TAAATCTACCGTGGACC 375
Db 6469 CAAATCGTTCTTCAAAATGGAGCCCCAGATGAGTCCATGACTTAAAGATCTTACCATGGACC 6528
Qy 376 CCTGGACCGGCTGCTAGATGATGCTCTGATGTTAATGACATTTGAAGTACCCCTCCCGA 435

Db 6529 CCTGGACCGGCTGCTAGCCATGCTGATGTTAATGACATCGAAGACACCCCTCCAGA 6588
Qy 436 GGAATATCAACTGTCACAACCCCTACTACTACTCACTCAATTCAGTAGGAAGCAGTTAGAGCAG 495
Db 6589 GGAATATCAACTGTCACAACCCCTACTACTATGCCCTGATTCAGCAGGAAGTAGTTAGAGTGG 6648
Qy 496 TTGTCAGCCCAACTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGA 555
Db 6649 TTGTCGGCCCAACTCCCAACAGCAGCTTGGTGGTCTGTTGGGAGGGGAGCTGAGAGA 6708
Qy 556 CAGGACTAGCTGATTTCTCTA-----GGCTGACTAGAAATCCCNAAAGCCTTANCTGGGA 608
Db 6709 CAGGACTAGCTGGGATTTCCCTAGGGCCGACTTAAGAATCAACACTAAGCGCTAGCTGGGA 6768
Qy 609 AGGTGACCGCAT-CCATCTTTTAAACATGGGGCTTGCAACTTAGCTCACAACCCGACCAATC 667
Db 6769 AGGTGACCGCGTCCCACTTTTAAACACGGGGCTTGCAACTTAGCTCACAACCCGACCAATC 6828
Qy 668 -----AGAGAGCTCACTAAATGCTAATCAGSCAAAAAAGAGAGGTAAAGCAATAG 718
Db 6829 AGGTAGCAAGAGAGAGCTCGTTAAATGATAATTAGSCAAAAACAGGAGGTAAAGAAATAG 6888
Qy 719 CCAATCATC-TAATGCTGAGAGCAGCGGGAAGGAGCAAGGATTTGGGATATAAATCTCAG 777
Db 6889 CCAATCATCTTATCAGCTGAGAGCAATAGGGAAGGAGCAATGATCGGGATATAAATCCAG 6948
Qy 778 GCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATCTGTTATGGGAGCTCTGT 837
Db 6949 GCATTCCAGTTGGCAACTGCTACCTCTTTGGGTCCTCCCTCTTGTATGGAAGCTCTGT 7008
Qy 838 TTTCACTCTATTTCACCTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTTCTTTTA 897
Db 7009 -----TTTCACTCTATTAAATCTTCTGCACTCTTCTGGTCCATATTGTTA 7058
Qy 898 TGCTCAAGCTGAGCTTTTGTGCGCATCCACACTGCTGTTGTCACCGCTCAGACACC 957
Db 7059 CGGCTCCAGCTGAGCTTTTGTGCGCGTCCACCCCTGTTGTTGCGCGCTGTCGACAGACC 7118
Qy 958 GCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGCTGCTCTGATCCAGCGA 1017
Db 7119 GCGCTGACTTCCATCCCTCCCGATCCAGAGGTTGTCGCTGCTCTGATGACGCGA 7178
Qy 1018 GGTACCCATTTGCCACTCCCGATCAGGCTAAAGGCTTGCATTTCTCTGATGGCTAAGT 1077
Db 7179 GGGCCCATTTCCACTCTGATCGGACTGAGGCTTGCCATTTCTCTGACGCTAAGA 7238
Qy 1078 GCCTGGGTTTGTCTTAATAGAACTGAACACTGTGTCTGTTGTTCTGTTCTTCTTCCAT 1137
Db 7239 GCCTGGGTTTGTCTTAATCGAGCTGAACACTAGTCACTGGGTTTCCAGGGTTCTTCTCCGT 7298
Qy 1138 GACCCAGGCTTCTAATAGAGCTAATCACTACCGCATGCGCCCAAGATTTCCATTTCTTCTG 1197
Db 7299 GACCCAGGCTTCTAATAGAGCTAATCACTACCGCATGCGCCCAAGATTTCCATTTCTTCTG 7358
Qy 1198 GTATCTGTAGGCGCAAGAACCCCGAGTTCAGAGAAAGTGAAGGCTTGCACCATTTGGGAAG 1257
Db 7359 GAATCGTGTAGGCGCAAGAACCCCGAGTTCAGAGAAACAGAGGCTTGCACCGCTCTTGAAG 7418
Qy 1258 TGSCCCACTGCCCCATTTTGTGAGCGGCCCAACCATCTTTGGGAGCTGTGGGAGCAAGAT 1317
Db 7419 TGSCCTGCCACCATTTGGAAGGGGCTGCCACCATCTTTGGGAGCTGTGGGAGCAAGGAC 7478
Qy 1318 CCCCAGTAACA 1329
Db 7479 CCCCAGTAACA 7490
RESULT 13
AAD41225
ID AAD41225 standard; cDNA; 2074 BP.
XX
XX AAD41225;
XX

XX 30-OCT-2002 (first entry)
 XX Human EMBRY-2 cDNA.
 XX
 KW Human; embryogenesis associated protein; AIDS; reproductive disorder;
 KW infertility; endometriosis; endometrial tumour; inflammatory disorder;
 KW autoimmune disorder; acquired immune deficiency syndrome; transgenic;
 KW ovarian tumour; contact dermatitis; placenta disorder; preclampsia;
 KW EMBRY-2; allergy; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 XX CDS 64..1692
 FT /*tag= a
 FT /product= "EMBRY-2 protein"
 FT sig_peptide 64..123
 FT /*tag= b
 FT mat_peptide 124..1689
 FT /*tag= c
 FT /product= "Mature EMBRY-2 protein"
 XX
 XX WO200248362-A2.
 XX
 XX 20-JUN-2002.
 XX
 XX 14-NOV-2001; 2001WO-US043956.
 XX
 XX 15-NOV-2000; 2000US-0249407P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Ramkumar J, Arvizu C;
 XX
 XX WPI; 2002-537629/57.
 XX P-PSDB; AAE25054.
 XX
 XX New polypeptides of human embryogenesis associated proteins for screening
 XX modulators useful for treating or preventing disorders e.g.
 XX endometriosis, infertility, allergy, preclampsia.
 XX
 XX Claim 59; Page 96-97; 97pp; English.
 XX
 XX The invention relates to human embryogenesis associated proteins (EMBRY)
 XX and nucleic acid molecules encoding such proteins. EMBRY sequences are
 XX useful for screening modulators useful for treating or preventing
 XX disorders associated with abnormal expression of EMBRY. The disorders
 XX treated include reproductive disorders such as infertility,
 XX endometriosis, endometrial or ovarian tumour; autoimmune/inflammatory
 XX disorder such as acquired immune deficiency syndrome (AIDS), allergies,
 XX contact dermatitis; disorders of the placenta such as preclampsia,
 XX abruptio placentae etc. Sequences of the invention are also useful for
 XX analysing a proteome of a tissue or a cell type. EMBRY proteins are
 XX useful as immunogens for preparing antibodies. Polynucleotides of the
 XX invention are useful for creating knockin humanised animals or transgenic
 XX animals to model human diseases. They are also used in gene therapy. The
 XX present sequence is human EMBRY-2 cDNA
 XX
 XX Sequence 2074 BP; 583 A; 567 C; 392 G; 531 T; 0 U; 1 Other;
 XX
 XX Query Match 57.7%; Score 766.8; DB 6; Length 2074;
 XX Best Local Similarity 93.0%; Pred. No. 1.4e-240;
 XX Matches 812; Conservative 0; Mismatches 60; Indels 1; Gaps 1;
 XX
 XX 1 TCAGAAATCGAAGAGCTTTAGACTTCTAACCAGCCCAAGAGAGGGGAACTGTTATTTT 60
 XX |
 XX 1203 TCAGAAATCGAAGAGCTTTAGACTTCTAACCAGCCCAAGAGAGGGGAACTGTTATTTT 1262
 XX |
 XX 61 AGGGAAGAAGTCTGTTAGTATGTTAAATCAATCTCGAATCAATCTGAGAAAGTTAAAGA 120
 XX |
 XX 1263 AGGGAAGAAGTCTGTTATGTTAAATCAATCAATCCAGAAATGCTGCTAGGAAAGTTAAAGA 1322
 XX |

QY 121 AATTGAGATCGAATAATAATGTAGACAGAGAGCTTTCAAAACACTGCACCTCGGGGCT 180
 DB |
 DB 1323 AATTCGAGATCGAATAATAATGTAGACAGAGAGAGCTTTCAAAACACTGCACCTCGGGGCT 1382
 QY 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTTCTTAGGACCTTAGCAGCTATAATATT 240
 DB |
 DB 1383 CCTCAGCAATGGATGCCCTGGGTTCTCCCTTTCTTAGGACCTTAGCAGCTATAATATT 1442
 QY 241 TTACTCTCTTTGGAGCCCTGTATCTCAACTTCTTTGAAGTTTCTCTTCCAGAAAT 300
 DB |
 DB 1443 GTTACTCTCTTTGGAGCCCTGTATCTTTAACTCTCTTTGAAGTTTCTCTTCCAGAAAT 1502
 QY 301 TGAAGCTGTAAGCTACAAATAGTCTTCAATGGAACCCAGATGAGTCCATGACTAA 360
 DB |
 DB 1503 TGAAGCTGTAAGCTACAAATAGTCTTCAATGGAACCCAGATGAGTCCATGACTAA 1562
 QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACTGA 420
 DB |
 DB 1563 AATCTACCGAGACCCCTGGACCGGCTGTAGCCATGCTCCGATGTTAATGACTGA 1622
 QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCACACCCCTACTACACTCCCAATTCAGTAGG 480
 DB |
 DB 1623 AGGCATCTCTCCCAAGGAAATCTCAACTGCACACCCCTACTATGCCCAATTCAGTAGG 1682
 QY 481 AAGCAGTTAGAGCAGTTGTGAGCCAACTCTCCCAACAGTACTTGGGTTTTCTCTGTGAGA 540
 DB |
 DB 1683 AAGCAGTTAGAGCAGTTGTGAGTCACTCTCCCAACAGCAGTCTGGGTTTTCTCTGTGAGA 1742
 QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCCAAGGCT 600
 DB |
 DB 1743 GGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCCGATTAAGAAATCCCTAAGGCT 1802
 QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
 DB |
 DB 1803 AGCTGGGAAGGTGACCGCGTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCG 1862
 QY 661 ACCAATCAGAGAGCTCACTAAATGCTAAATCAGGCAAAAACAGGAGGTAAAGCAATAGCC 720
 DB |
 DB 1863 ACCAATCAGAGAGCTCACTAAATGCTAAATGCTAAATGCTAAATGCTAAATGCTAAATAGCC 1922
 QY 721 AATCATCTATTGCTGAGAGACAGCGGGAAGGCAAGGATTTGGGATATAAATCAGGCA 780
 DB |
 DB 1923 AATCATCTATTGCTGAGAGACAGTGGGAGGGAAGGATTTGCAATATAAACCAGGCA 1982
 QY 781 TTCAAGCCCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
 DB |
 DB 1983 TTCAGGCCAGC-ANAGCAACCGCTTTGGGTCCCTTCCCTTGTATGGAGCTCTGTTTT 2041
 QY 841 CACTCTATTTTCACTCTATTAAATCATCAACTG 873
 DB |
 DB 2042 CACTCTATTTCACTCTATTAAATCTTGCAACTG 2074
 XX
 XX RESULT 14
 XX ACC46747
 XX ID ACC46747 standard; cDNA; 2046 BP.
 XX AC ACC46747;
 XX XX
 XX 02-JUN-2003 (first entry)
 XX
 XX Human dithp growth/development-associated protein-encoding cDNA.
 XX
 XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 XX cancer; cell proliferative disorder; autoimmune disorder;
 XX inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 XX neurological disorder; gastrointestinal disorder; transport disorder;
 XX connective tissue disorder; drug screening; proteome analysis;
 XX gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 XX disease model; toxicological testing; transcript imaging; growth;
 XX development; gene; ss.
 XX
 XX Homo sapiens.

ID AAX77526 standard; cDNA; 2946 BP.
XX
AC AAX77526;
XX
DT 10-AUG-1999 (first entry)
XX
DE Human secreted protein AJ172_2 cDNA.
XX
KW Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW cell proliferation; cell differentiation; suppressor; tumour inhibitor;
KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumour;
KW cadherin; tumour invasion suppressor; gene therapy; tissue growth; ss.
XX
OS Homo sapiens.
XX
PN WO9926972-A1.
XX
PD 03-JUN-1999.
XX
PF 17-NOV-1998; 98WO-US024614.
XX
PR 21-NOV-1997; 97US-00976110.
PR 18-MAY-1998; 98US-00080478.
PR 20-OCT-1998; 98US-00175928.
XX
XX (GENY) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M;
XX
XX WPI; 1999-357813/30.
XX P-PSDB; AAY08622.
XX
PT New polynucleotides encoding secreted proteins.
XX
XX Claim 13a; Page 100-101; 142pp; English.
XX
CC This invention describes novel human secreted proteins encoded by
CC polynucleotides isolated from human adult testes, adult brain, adult
CC blood or adult placenta, or murine adult bone marrow or thymus cDNA
CC libraries. The products of the invention are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals, although no
CC supporting data is given. Suggested activities include nutritional
CC activity, cytokine and cell proliferation/differentiation activity,
CC immune stimulating (e.g. as vaccines) or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, and tumour
CC inhibition activity. The polynucleotides are also stated to be useful for
CC gene therapy
XX
SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;
Query Match 55.1%; Score 731.8; DB 2; Length 2946;
Best Local Similarity 90.7%; Pred. No. 5.7e-229;
Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAGAGGGGGAACCTGTTATTTT 60
DB 2067 TCAAAATCGAAGAGCTTTAGACTTCTTAACCGCTGAAAGAGGGGGAACCTGTTATTTT 2126
QY 61 AGGGGAAGAAATCGCTGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAGTTAAAGA 120
DB 2127 AGGGGAAGAAATCGCTGTATTTATTTAATCAATCCGAATCGTCACTGAGAAGTTAAAGA 2186
QY 121 AATTTGAGATCGAATATAATGATAGACGAGAGACCTTCAAAACACTGCACCCCTGGGGCT 180
DB 2187 AATTCGAGATCGAATACACGATAGACGAGAGAGCTTCCAAACACTGGACCCCTGGGGCT 2246

QY 181 CCTCAGCCAATGGATGCCCTGGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 240
DB 2247 CCTCAGCCAATGGATGCCCTGGATCTCCCTCTTAGGACCTCTAGCAGCTATAATATT 2306
QY 241 TTTACTCTCTTTTGGACCCCTGTATCTTCAACTCTCTTTGTTAAGTTTGTCTTCCAGAAAT 300
DB 2307 GCTACTCTCTTTGGACCCCTGTATCTTAACTCTCTTTGTTAAGTTTGTCTTCCAGAAAT 2366
QY 301 TGAAGCTGTAAAGCTACAATATAGTTCTTCAATAGGAACCCAGATGAGTCCATGACTAA 360
DB 2367 CGAAGCTGTAAACTA-----CAAATGGAGCCCAAGATGCAAGTCCCAAGACTAA 2414
QY 361 AATCTACGCTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 2415 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCGAGATCTGATGTTAATGACATCAA 2474
QY 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCAACAAACCCCTACTACATCTCAATTCAGTAGG 480
DB 2475 AGGCACCCCTCCTGAGGAAATCTCAGCTGCAACACCTCTACTAGCCCCCAATTCAGCAGG 2534
QY 481 AAGCAGTTAGAGCAGTTGTGAGCCAACTCCCAACAGTACTTGGGTTTCTCTTTGAGA 540
DB 2535 AAGCAGTTAGAGCAGTCTGTCGGCCAACTCCCAACAGCAGTACTTGGGTTTCTCTTTGAGA 2594
QY 541 GGGTGGACTGAGACAGGAGCTAGCTGGATTTCTTAGGCTGACTAAGAATCCCAAGGCT 600
DB 2595 TGGGGGACTGAGACAGGAGCTAGCTGGATTTCTTAGGCTGACTAAGAATCCCTAAGGCT 2654
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCGG 660
DB 2655 AGCTGGGAAGGTGACCATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2714
QY 661 ACCAATCAGAGAGCTCAATAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
DB 2715 ACCAATCAGAGAGCTCAATAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCC 2774
QY 721 AATCATCTATTGCTGAGACGACGAGGAAAGCAAGGATTGGGATATAAACTCAGGCA 780
DB 2775 AATCATCTATTGCTGAGACGACGAGGAGGACATGATCGGGATATAAACCCAGTC 2834
QY 781 TTCAAGCCGACCAACAGCAACCCCTTTGGGTCCTCCCTCCATGTATGGGAGCTCTGTTTT 840
DB 2835 TTCAGCGCGCAACGGCAACCCCTTTGGGTCCTCCCTTTGTATGGGAGCTCTGTTTT 2894
QY 841 CACTCTATTTTCACTCTATTAAATCATGCAACTGCA 875
DB 2895 CATGCTATTTTCACTCTATTAAATCTTGAOACTGCA 2929
RESULT 16
AAZ59468
ID AAZ59468 standard; cDNA; 2946 BP.
XX
AC AAZ59468;
XX
DT 11-APR-2000 (first entry)
XX
DE Human secreted protein AJ172_2 polynucleotide sequence.
XX
KW Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
KW placental pathology; metastasis inhibition; nutritional activity;
KW immune stimulator; haematopoiesis regulator; tissue growth;
KW tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9960020-A1.
XX
PD 25-NOV-1999.
XX
PF 17-MAY-1999; 99WO-US010915.
XX

PR 18-MAY-1998: 98US-00080478.
PR 20-OCT-1998: 98US-00175928.
XX (GEMY) GENETICS INST INC.
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Mi S, Treacy M;
XX WPI; 2000-116311/10.
DR P-PSDB; AAY67313.
XX
XX New polynucleotides encoding secreted cDNA libraries, used to develop
PT products for the diagnosis and treatment of neoplastic disease.
XX
XX Claim 14; Page 107-108; 149pp; English.
XX This is the human secreted protein AJ172.2 nucleotide sequence, obtained
CC from a human adult testes cDNA library. The invention relates to secreted
CC human and murine proteins. The polynucleotides and proteins are predicted
CC to have biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals. Detection of the levels of the proteins can be used for the
CC diagnosis of e.g. pre-eclampsia, placental pathology or cancer. Agents
CC which modulate the expression or function of the proteins may be used for
CC treating a neoplastic disease and inhibiting metastasis. Other suggested
CC activities include nutritional activity (e.g. in feeds), cytokine and
CC cell proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity.
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotide sequences are also stated to be useful for gene therapy
XX
XX Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;

Query Match 55.1%; Score 731.8; DB 3; Length 2946;
Best Local Similarity 90.7%; Pred. No. 5.7e-229;
Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTTATTTT 60
DB 2067 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTT 2126
QY 61 AGGGGAAGATGCTGTTAGTATGTTTAACTGAATCTGAATCATTTACTGAGAACTTAAAGA 120
DB 2127 AGGGGAAGATGCTGTTATTTATTTATTTAATCAATCCGAATCGTCACTGAGAAAGTTAAAGA 2186
QY 121 AATTTGAGATCGAATATAATGTAGAGCAGAGGACCTTCAAACACTGCACCTGGGCGCT 180
DB 2187 AATTCGAGATCGAATACAAGTGAAGAGAGGAGCTTCGAAACACTGGACCTGGGCGCT 2246
QY 181 CCTCAGCAATGGATGCGCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATT 240
DB 2247 CCTCAGCAATGGATGCGCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATATT 2306
QY 241 TTTACTCTCTTTGGACCTGTTATCTCAACTCTCTTTTAACTTTGTTCTCTCCAGAA 300
DB 2307 GCTACTCTCTTTGGACCTGTTATCTTTAACTCTCTTTGTTAACTTTGTTCTCTCCAGAA 2366
QY 301 TGAAGCTGTAAGACTACAAATAGTTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAA 360
DB 2367 CGAAGCTGTAAGACTA-----CAATGGAGCCCAAGATGCGTCCAGACTAA 2414
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTAGTCTGATGTTTAAATGACATTTGA 420
DB 2415 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTTAAATGACATCAA 2474
QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGACAAACCCCTACTACCTCCAAATTCAGTAGG 480
DB 2475 AGGCACCCCTCCTGAGGAATCTCAGCTGCACAACTCTACTAGCCGCCAAATTCAGCAGG 2534
QY 481 AAGCAGTTAGCAGTGTGTCAGCAACCTCCCCAACAGTACTTGGGTTTTCTCTGTTGAGA 540

DB 2535 AAGCAGTTAGAGCGTCTGCGCCAAACCTCCCAACAGCAGCTTAGCTTTCTGTTGAGA 2594
QY 541 GGGTGGACTGAGAGCAGAGCTAGCTGGATTTCTTAGCTAGCTAAGAAATCCNAAAGCCT 600
DB 2595 TGGGGGACTGAGAGCAGAGCTAGCTGGATTTCTTAGCTAGCTAAGAAATCCCTAAGCCT 2654
QY 601 ANCTGGGAAGGTGACCGCATCTTTAAACATGGGGCTTCGCAACTTAGCTACACCCG 660
DB 2655 AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGCTACACCTG 2714
QY 661 ACCAATCAGAGAGCTCACTAAATCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
DB 2715 ACCAATCAGAGAGCTCACTAAATCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2774
QY 721 RATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTTGGGATATATAAATCTCAGGCA 780
DB 2775 AATCATCTATTGCTGAGAGCAGCGGAGGAGCAATGATCGGATATATAAATCAGGCA 2834
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCACTTGTATGGAGCTCTGTTTT 840
DB 2835 TTGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTT 2894
QY 841 CACTCTATTTCATCTCTATTAAATCATGCAACTGCA 875
DB 2895 CATGCTATTTCATCTCTATTAAATCTTGCAACTGCA 2929

RESULT 17
ADC38776
ID ADC38776 standard; cDNA; 2946 BP.
XX
XX ADC38776;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human cDNA encoding a secreted protein #63.
XX
XX ss; gene; immune disorder; severe combined immunodeficiency; SCID;
XX autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
XX rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;
XX lymphoid cell deficiency; osteoporosis; osteoarthritis;
XX peripheral nervous system disease; peripheral neuropathy;
XX Alzheimer's disease; Parkinson's disease; coagulation disorder;
XX inflammatory disease; systemic inflammatory response syndrome; SIRS;
XX ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;
XX hypersensitivity; regeneration; neural cell proliferation; fertility;
XX tumour; chemokine; human; secreted protein.
XX
XX Homo sapiens.
XX
XX US2002193567-A1.
XX
XX 19-DEC-2002.
XX
XX 02-APR-2002; 2002US-00114893.
XX
XX 11-AUG-1995; 95US-00514014.
XX 05-APR-1996; 96US-00628364.
XX 19-APR-1996; 96US-00635311.
XX 07-JUN-1996; 96US-00659224.
XX 17-JUN-1996; 96US-00664596.
XX 09-JUL-1996; 96US-00677231.
XX 26-JUL-1996; 96US-00686878.
XX 23-AUG-1996; 96US-00701819.
XX 27-SEP-1996; 96US-00721488.
XX 27-SEP-1996; 96US-00721798.
XX 27-SEP-1996; 96US-00721923.
XX 27-SEP-1996; 96US-00721926.
XX 25-OCT-1996; 96US-00738367.
XX 30-OCT-1996; 96US-00739775.
XX 13-JAN-1997; 97US-00783395.
XX 10-APR-1997; 97US-00833823.

PR 02-JUN-1997; 97US-00867677.
 PR 05-SEP-1997; 97US-00924838.
 PR 06-OCT-1999; 99US-00413232.
 XX (GENY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;
 PI Kelleher K;
 XX
 DR WPI; 2003-657236/62.
 DR P-PSDB; ADC38777.
 XX
 PT Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and
 PT BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,
 PT useful for treating e.g. multiple sclerosis and rheumatoid arthritis.
 XX
 PS Disclosure; SEQ ID NO 134; 412pp; English.
 XX
 CC The invention relates to a protein comprising fully defined AZ302 1
 CC protein or BD127 1 6 protein. The polynucleotides are useful for
 CC expressing recombinant proteins for analysis and are also useful as
 CC chromosome markers or tags to identify chromosomes or to map related gene
 CC positions. The proteins are useful as amino acid supplement. Carbon
 CC source, nitrogen source and carbohydrate source. The proteins are useful
 CC for treating various immune deficiencies and disorders (e.g. severe
 CC combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic
 CC reactions (e.g. asthma), myeloid or lymphoid cell deficiencies,
 CC osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.
 CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),
 CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory
 CC response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease),
 CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing
 CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or
 CC nerve growth or regeneration, for proliferating neural cells and for
 CC regenerating nerve and brain tissue, for inducing fertility and for
 CC inhibiting tumour growth. Proteins are also useful as chemokine for
 CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also
 CC useful as inhibitors of receptor/ligand interactions. The present
 CC sequence represents cDNA encoding a human secreted protein.
 XX
 SQ Sequence 2946 BP; 858 A; 801 C; 570 G; 717 T; 0 U; 0 Other;
 Query Match 55.1%; Score 731.8; DB 10; Length 2946;
 Best Local Similarity 90.7%; Pred. No. 5.7e-229;
 Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
 QY 1 TCARAATCGAAGAGCTTTAGACTTGCTAACCGCCMAAGAGGGGGAACCTGTTTATTTT 60
 DB 2067 TCARAATCGAAGAGCTTTAGACTTGCTAACCGCTGNAAGGGGGGAACTGTTTATTTT 2126
 QY 61 AGGGGAAGAATGCTGTTAGTATGTTTAAATCAATCTGGAATCATTCTGAGAAAAGTTAAAGA 120
 DB 2127 AGGGGAAGAATGCTGTTTATTTATTTATTTAAATCAATCCGGAATCGTCACTGAGAAAAGTTAAAGA 2186
 QY 121 AATTGAGATCGAATATATGAGACGAGGACCTTCAAAACACTGACCCCTGGGGCT 180
 DB 2187 AATTGAGATCGAATCAACGTAAGACGAGGAGCTTCAAAACACTGAGACCTGGGGCT 2246
 QY 181 CCTCAGCAATGATGCTGCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
 DB 2247 CCTCAGCAATGATGCTGCTGGATTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 2306
 QY 241 TTTTACTCTCTTTGGACCCCTGATCTTCAACTCTCTTGTAGTTTGTCTCTCCAGAT 300
 DB 2307 GCTACTCTCTTTGGACCCCTGATCTTAACTCTCTTGTAACTTTGTCTCTTCCAGAT 2366
 QY 301 TGAAGCTGTAAGCTCAAAATAGTTCTTCAATGGAAACCCAGATGAGTCCATGACTAA 360
 DB 2367 CGAAGCTGTAAACTA-----CAATGGAGGCCCAAGATGAGTCCCAAGACTAA 2414
 QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420

DB 2415 GATCTACCGCAGAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 2474
 QY 421 AGTCACCCCTCCGAGAGAAATCTCAACTGCGACAAACCCCTACTACTACCTCCAAATTCAGTAGG 480
 DB 2475 AGGCACCCCTCCTGAGGAAATCTCAGCTGCGACAAACCTCTACTACGCCCCCAATTCAGCAGG 2534
 QY 481 AAGCAGCTTAGAGCAGTGTGTCAGCAACCTCCCAACAGTACTTTGGGTTTCTCTGTTGAGA 540
 DB 2535 AAGCAGTTAGAGCGGCTGTCGGCCAACTCCCAACAGCAGCACTTAGGTTTCTCTGTTGAGA 2594
 QY 541 GCGTGGACTGAGACAGCAGTACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCAAGCCT 600
 DB 2595 TGGGGGACTGAGACAGCAGTACTAGCTGGATTTCCTAGGCTGACTAAGAATCCCTAAGCCT 2654
 QY 601 ANCTGGGAAGGTGACCGCATCTTTAAACATGGGGCTTGCAACTTAGCTACACACCCG 660
 DB 2655 AGCTGGGAAGGTGACCATCTCCACCTTTAAACACGGGGCTTGCAACTTAGCTACACCTG 2714
 QY 661 ACCAATCAGAGAGCTCACTAAAATGCTAAATCAGGCAAAAAACAGAGGTAAGCAATAGCC 720
 DB 2715 ACCAATCAGAGAGCTCACTAAAATGCTAAATCAGGCAAAAAACAGAGGTAAGCAATAGCC 2774
 QY 721 AATCATCTATTGCTGAGACAGCAGCGGGAAGGACAAGGATTGGGATATAAATCTCAGCA 780
 DB 2775 AATCATCTATTGCTGAGACAGCAGCGGGAAGGACAATGATCGGGATATAAACCCTAAGTC 2834
 QY 781 TTCAGGCGACACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGGAGCTCTGTTTT 840
 DB 2835 TTGAGCGGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTT 2894
 QY 841 CACTCTATTTCACTCTTATTAAATCATGCAACTGCA 875
 DB 2895 CATGCTATTTCACTCTTATTAAATCTTCAACTGCA 2929
 RESULT 18
 AAD24195
 ID AAD24195 standard; cDNA; 2930 BP.
 AC AAD24195;
 XX
 DT 07-MAY-2002 (first entry)
 XX
 DE Human syncytin cDNA.
 XX
 KW Human; syncytin; preeclampsia; gestational trophoblast disorder;
 KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;
 KW envelope gene; human endogenous defective retrovirus; HERV-W; 88.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 930..2546
 FT /*tag= a
 FT /product= "Syncytin"
 XX
 WO200204678-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 09-JUL-2001; 2001WO-US021719.
 XX
 PR 07-JUL-2000; 2000US-0216657P.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Keith JC, McCoy JM, Mi S;
 XX WPI; 2002-171727/22.
 DR P-PSDB; AAE14540.
 XX
 PT Identifying a compound for treating a subject with or at risk of

PT developing preclampsia, comprises determining whether the expression or
PT activity of syncytin in the cell is modulated in the presence of a test
XX compound.

XX Disclosure; Page 39-42; 43pp; English.

PS The invention relates to identifying compounds which are modulators of
XX syncytin expression. The syncytin modulators are useful in diagnosis and
XX treatment of preclampsia and gestational trophoblast disorders (e.g.
XX choriocarcinoma, hydatiform mole, placental site tumour and missed/
XX incomplete abortion). Syncytin is a human gene derived from the envelope
XX gene of human endogenous defective retrovirus, HERV-W. The present
XX invention is based partly on the discovery that syncytin expression is
XX dramatically reduced in preclampsia, and is also mis-localised to the
XX apical syncytiotrophoblast membrane. The present sequence is human
XX syncytin cDNA

SQ Sequence 2930 BP; 842 A; 800 C; 571 G; 717 T; 0 U; 0 Other;

Query Match 55.0%; Score 730.8; DB 6; Length 2930;
Best Local Similarity 90.7%; Pred. No. 1.2e-228;
Matches 793; Conservative 0; Mismatches 69; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCCMAAAGAGGGGAACTGTTTATTTT 60
DB TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAAGAGGGGAACTGTTTATTTT 2128

QY 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAATTAAGA 120
DB AGGGGAAGAAATGCTGTATTATGTTAATCAATCCGGAATCGTCACTGAGAAATTAAGA 2188

QY 121 AATTCGAGATCGAATATAATAGACAGAGGACCTTCAAAACACTGCACTTGGGGCT 180
DB AATTCGAGATCGAATACAACGTAGACAGAGGAGCTTCGAAACACTGGAACCTTGGGGCT 2248

QY 181 CCTCAGCAATGATGCTCCCTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
DB CCTCAGCAATGATGCTCCCTGGATTTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 2308

QY 241 TTTACTCTCTTTGGACCCCTGATCTCAACTCTCTTGTGTTAGTTTGTCTTCTCCAGAT 300
DB GCTACTCTCTTTGGACCCCTGATCTTAACTCTCTTGTGTTAGTTTGTCTTCTCCAGAT 2368

QY 301 TGAAGCTGTAAAGCTACAATAGTTCTTCAATGGAAACCCAGATCGAGTCCATGACTAA 360
DB CGAAGCTGTAAAGCTA-----CAATGGAGCCCAAGATCGAGTCCAGACTAA 2416

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
DB GATCTACCGGAGACCCCTGGACCGGCTGCTAGCCCAAGTCTGATGTTAATGACATCAA 2476

QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTACATCCCAATTCAGTAGG 480
DB AGGCACCCCTCCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGG 2536

QY 481 AAGCAGTTAGACAGTTGTCAGCAACCTCCCAACAGTACTTGGTTTCTCTGTTGAGA 540
DB AAGCAGTTAGAGGGTGTGTCGGCAACCTCCCAACAGCACTTAGGTTTCTCTGTTGAGA 2596

QY 541 GGGTGGACTCAGACAGAGGACTAGTGGATTTCTTAGGCTGACTTAAGAAATCCCNAAAGCCT 600
DB TGGGGGACTCAGACAGAGGACTAGTGGATTTCTTAGGCTGACTTAAGAAATCCCTAGCCT 2656

QY 601 ANCTGGGAAGTGACCCGATCCATCTTTTAAACATGGGGCTTGGCACTTAGCTCACACCCG 660
DB AGCTGGGAAGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2716

QY 661 ACCAATCAGAGAGCTCTAATAATGCTTAATCAGGCAAAACAGAGGTAAGCAATAGCC 720
DB ACCAATCAGAGAGCTCTAATAATGCTTAATAGGCAAAACAGAGGTAAGCAATAGCC 2776

QY 721 AATCATCTATTGCTGAGACACAGCGGGAGGACAGAGGATTTGGGATATATAAATCTCAGGCA 780

DB 2777 AATCATCTATTGCTGAGAGCACAGAGGAGCAATGATCGGATATAAACCAAGTC 2836

QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTTT 840
DB TTCAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCATTTGATGGAGCTCTGTTTT 2896

QY 841 CACTCTATTTTCACTCTATTAAATCATGCAACTGC 874
DB CATGCTATTTCACCTCTATTAAATCTTGCAACTGC 2930

RESULT 19
AAF55630
ID AAF55630 standard; DNA; 2781 BP.
XX AAF55630;
XX
XX 29-MAY-2001 (first entry)
XX Nucleotide sequence of a human endogenous retrovirus envelope protein.
XX Envelope protein; HERV; syncytia formation; placental development;
XX syncytia; cancer; cell adhesion; ss.
XX Human endogenous retrovirus.
XX
XX Key Location/Qualifiers
XX CDS 762..2378
XX FT /*tag= a
XX FT /product= "envelope protein"
XX
XX WO200116171-A1.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-FR002429.
XX
XX 01-SEP-1999; 99FR-00011141.
XX 15-SEP-1999; 99FR-00011793.
XX
XX (INMR) BIO MERIEUX.
XX (INRM) INST NAT SANTE & RECH MEDICALE.
XX
XX Mallet F, Cosset F, Blond J, Lavillette D, Bouton O, Ruggieri A;
XX WPI; 2001-226676/23.
XX P-PSDB; AAB67652.
XX
XX Detecting expression of human endogenous retrovirus envelope protein in
XX cells of a tissue or culture, from its ability to induce syncytia.
XX
XX Disclosure; Page 44-45; 57pp; French.
XX
XX The present sequence encodes a human endogenous retrovirus envelope
XX protein. The specification describes a method for detecting expression of
XX an envelope protein from a human endogenous retrovirus (HERV), in cells,
XX of a tissue or culture. The method comprises detecting syncytia formation
XX due to the fusogenic properties of the envelope protein. Envelope
XX polypeptides and polynucleotides are used to produce therapeutic or
XX prophylactic compositions, particularly for treatment of cancer, to
XX correct defects in placental development (or other natural formation of
XX other types of syncytia), and to promote adhesion of cells in grafts or
XX cellular repair processes. Expression of sequences antisense to the
XX polynucleotide are used to prevent formation of syncytia
XX
XX
XX Sequence 2781 BP; 747 A; 770 C; 556 G; 708 T; 0 U; 0 Other;

Query Match 54.9%; Score 730.2; DB 5; Length 2781;
Best Local Similarity 90.6%; Pred. No. 1.8e-228;
Matches 793; Conservative 0; Mismatches 70; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTTTATTTT 60

Db 1901 TCAAAATCGAAGAGCTTTAGACTTGCTAAACGCTGAAGAGGGGAACTGTTATTTTT 1960
QY 61 AGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTAAGTGAAGAAGTTAAAGA 120
Db 1961 AGGGAAGAATGCTGTTATGTTAATCAATCGGAATCGTCACTGAGAAGTTAAAGA 2020
QY 121 AATTGAGATCGAATATATATAGTAGAGAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 2021 AATTGAGATCGAATATCAACGCTAGAGAGAGAGGAGCTTCGNAACACTGCACCTGGGGCT 2080
QY 181 CCTCAGCAATGAGTGGCCCTGAGCTCTCCCTTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 2081 CCTCAGCAATGAGTGGCCCTGAGTCTCCCTTCTTTAGGACCTCTAGCAGCTATAATATT 2140
QY 241 TTTACTCTCTTTTGGACCTGATCTTCAACTCTCTGTTAAGTTTGTCTCTTCAGAAAT 300
Db 2141 GCTACTCTCTTTGGACCTGATCTTAAACCTCTCTGTTAAGTTTGTCTCTTCAGAAAT 2200
QY 301 TGAAGCTGAAGCTACAATAGTCTTCAATGGAACCCAGATGCGAGTCCATGACTAA 360
Db 2201 CGAAGCTGAAGCTA-----CAATGGAGCCCAAGATGCGAGTCCAGACTAA 2248
QY 361 AATCTACCGTGGACCCCTGGAGCCGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2249 GATCTACCGCAGACCCCTGGAGCCGCTGTAGCCACAGATCGATGTTAATGACATCA 2308
QY 421 AGTACCCCTCCGAGGAAATCTCAATGCGACAAACCCCTATCACTCAATTCAGTAGG 480
Db 2309 AGGACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 2368
QY 481 AAGCAGTTAGACAGTCTGCGCCAACTCCCAACAGTCTTGGGTTTCTCTGTTGAGA 540
Db 2369 AAGCAGTTAGACAGTCTGCGCCAACTCCCAACAGTCTTGGGTTTCTCTGTTGAGA 2428
QY 541 GGGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 2429 TGGGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT 2488
QY 601 ANCTGGGAGGTGACCGATCCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCG 660
Db 2489 AGCTGGGAGGTGACCAATCCACTTTAAACACGGGGCTTTGCAACTTAGCTCACACCTG 2548
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2549 ACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCC 2608
QY 721 AATCATCTATTGCTGAGAGCACAGCGGAGGACAAGGATTTGGATATAAATCAGGCA 780
Db 2609 AATCATCTATTGCTGAGAGCACAGCGGAGGACAATGATCGGATATAAATCAGGCA 2668
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCTCCCATTTGATGGAGCTCTGTTT 840
Db 2669 TTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCTCCCATTTGATGGAGCTCTGTTT 2728
QY 841 CACTCTATTCTACTTATTAAATCATGCAACTGCA 875
Db 2729 CATGCTATTCTACTTATTAAATCTTGCAACTGCA 2763

RESULT 20
ID AAS84210 standard; cDNA; 6394 BP.
XX AC AAS84210;
XX 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #20014.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.

XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG20023.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 20014; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridization probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activities. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 6394 BP; 1840 A; 1597 C; 1384 G; 1571 T; 0 U; 2 Other;
Query Match 54.8%; Score 728.6; DB 5; Length 6394;
Best Local Similarity 78.6%; Pred. No. 1e-227;
Matches 1094; Conservative 0; Mismatches 223; Indels 74; Gaps 16;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGCTAAACGCTGAAGAGGGGAACTGTTATTTTT 60
Db 1175 TCAAAATCGAAGAGCTTTAGACTTGCTAAACGCTGAAGAGGGGAACTGTTATTTTT 1234
QY 61 AGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTAAGTGAAGAAGTTAAAGA 120
Db 1235 AGGGAAGAATGCTGTTATGTTAATCAATCGGAATCGTCACTGAGAAGTTAAAGA 1294
QY 121 AATTGAGATCGAATATATATAGTAGAGAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 1295 AATTGAGATCGAATATCAACCTAGAGCAGAGGAGCTTCGNAACACTGCACCTGGGGCT 1354
QY 181 CCTCAGCAATGAGTGGCCCTGAGCTCTCCCTTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 1355 CCTCAGCAATGAGTGGCCCTGAGTCTCCCTTCTTTAGGACCTCTAGCAGCTATAATATT 1414
QY 241 TTTACTCTCTTTTGGACCTGATCTTCAACTCTCTGTTAAGTTTGTCTCTTCAGAAAT 300
Db 1415 GCTACTCTCTTTGGACCTGATCTTAAACCTCTCTGTTAAGTTTGTCTCTTCAGAAAT 1474
QY 301 TGAAGCTGAAGCTACAATAGTCTTCAATGGAACCCAGATGCGAGTCCAGACTAA 360

1475 CGAGCTGTAAACTA-----CAATGGAGCCAGATGCACTCCAGACTAA 1522
1476
361 AATCTACCTGGACCCCT-GGACCGGCTGCTAGACTATGCTCTGATGTATGACATTTG 419
1477
1523 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTATGACATCA 1582
1478
420 AAGTCAACCTCCGAGGAAATCTCACTGCAACCCCTACTACACTCC-----469
1479
1583 AAGGCACCCCTCTGAGGAAATCTGAGCTGCAAAACCTCTACTTACGCCCAATTCAGG 1642
1480
470 -----AATTCAGTAGGAAGCAGTTAGAGCAGTGTGTGAGCCAACTCCCAACAGTACTTG 524
1481
1643 CAGGGAAGCAGTAGGAGCGGCTGTTGGGCCAACTTCCCAACAGCAGTCTTAGG 1702
1482
525 GGTTCCTGTTGAGAGGTGGACTGAGAG-----ACAGCAGTACTGGA--TTTCCTAGG 578
1483
1703 GTTTTCTGTTGGAGATGGGGGACTGAGGAGACAGGATTTAGCTGGGATTTCTCTAGG 1762
1484
579 CTGACTAAGNATCCCNAGCTANTCTGGGNAAGTGACCG-----CATCCATCT 626
1485
1763 CTGACCTAAGATCCCTAAGNCTTAGCTGGGGAAGGTGAGCAATCCACCTTTT 1822
1486
627 TTAACAATGGG--CTTGCACTTAGCTCACACCC--GACCAATCAGAGAGCTCACTA 680
1487
1823 TAAACACAGGGGGCTTTGCACTTTAGTTTACACTTGCACCAATCAGAGAGCTCACTA 1882
1488
681 AATG-CTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCTATTGCTTGA-G 738
1489
1883 AATGCTTAATAGGCAACAGAGGAGTAAAGAAATAGCAATCATCTATTGCTTGA 1942
1490
739 AGCAGAGCGGAAGGACAGGATTTGGATATAACTCAGGCATTCAGGCAGC-AACAGC 797
1491
1943 AGCAGAGTGGAGGCAAGGATTTGCAATATAAACCCAGGCAATTCAGGCAGGCAAGCC 2002
1492
798 AACCCCTTTGGGTCCCTCCCACTTGTATGGGAGCTCTGT-----TTTCACTCTA 847
1493
2003 AACCGCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGTTTCATGCTATTTCCTCTA 2062
1494
848 TTTCACTCTATAAATCATGCAACTG--CACTCTTCTGGTCCGTGTTTTTATGCTCAA 905
1495
2063 TTTCACTCTATAAATCTTCAACTGTACACTCTTCTGGTCACTGTTTATGAGCTCTA 2122
1496
906 GCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGCACCGTCA---CAGACCCGCTG 962
1497
2123 GCTGAGCTTCTGCTGAGTCCACCACTGCTGTTTGGCGCCGACCGCAGACCCGCTG 2182
1498
963 TGACTTCCATCCCTTTGGATCCAGAGAGTGTCCACTGTGCTCTGATCCAGCGAGTAC 1022
1499
2183 TGGCTTCCATCCCTTGGATCCGCGAGGTGTCTGCTGCATTCCTGATCCAGCGAGCGC 2242
1500
1023 CCATTGCACTCCCGATCAGGCTAAAGGCTTGCAATGTTCTGCAATGCTAGTGGCTG 1082
1501
2243 CCATTGCACTCCCGATCAGGCTAAAGGCTTGCAATGTTCTGCAATGCTAGTGGCTG 2302
1502
1083 GGTTCCTTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142
1503
2303 GGTTCCTTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2362
1504
1143 ACGCTTCTTAATAG--AGCTTAATACACTCACCGCATGGCCCAAGATTCATTCCTGCTGA 1200
1505
2363 ATGCTTCTTAATAGAACTATTAACTAACATCAGATGGGCCCAAGATTCATTCCTGCTGA 2422
1506
1201 --TCTGTAGGCCCAAGAACCCAGGTGAGAGANGTGAAGGCTTGCCACCAATTTGGGAGT 1258
1507
2423 ATTCGCTGAGGCCAAGAACCCAGGTGAGAGANGTGAAGGCTTGCCACCAATTTGGGAGT 2482
1508
1259 GGCCCACTGCCATTTGGTAGCGGCCCAACCACTTTGGGAGCTGTTGGGAGGAGGATC 1318
1509
2483 GGCCCGTGCCTATCTGGAAGCGGCTTGCCCAACCACTTTGGGAGCTGTTGGGAGGAGG 2542
1510
1319 CCCAGTAACA 1329
1511

Db 2543 CCCCGGTAAACA 2553
RESULT 21
AA25565
ID AAX25665 standard; cDNA to mRNA; 7582 BP.
XX
XX AAX25665;
XX
DT 21-MAY-1999 (first entry)
XX
DE Complete human endogenous retrovirus W genome.
XX
XX Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.
XX
PF 06-JUL-1998; 98WO-FR001442.
XX
PR 07-JUL-1997; 97PR-00008815.
XX
PA (INMR) BIO MERIEUX.
XX
PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX
XX WPI; 1999-120897/10.
DR
XX New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.
XX
XX Claim 1; Page 71-74; 106pp; French.
XX
CC This sequence represents the complete sequence of the human endogenous
CC retrovirus (HERV) W genome. The nucleic acids, their fragments or
CC peptides encoded by them are markers of autoimmune disease (e.g. multiple
CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
CC insulin-dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility
XX
XX Sequence 7582 BP; 2156 A; 1877 C; 1537 G; 1796 T; 2 U; 214 Other;
SQ
Query Match 54.4%; Score 722.6; DB 2; Length 7582;
Best Local Similarity 88.5%; Pred. No. 1.1e-225;
Matches 774; Conservative 19; Mismatches 70; Indels 12; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCGCAAAAGAGGGGAACTGTTATTTT 60
6720 TCEAATCGAAGAGCTTTAGACTTGTCTAACCGCTGAGAGAGGGGAACTGTTATTTT 6779
Db 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGCAATCATCTAGAGAAAGTAAAGA 120
6780 AGGGGAAGATGCTGTTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTAAAGA 6839
QY 121 AATTTCAGATCGAATATAATGTAGACAGAGGACCTTCAAAACACTGCCCTGGGSCCT 180
6840 AATTTCAGATCGAATATACTACAGTAKAGAGAGAGCTTCGAAACACTGGACCTGGGSCCT 6899
QY 181 CCTCAGCAATGATGCCCTGGACTCTCCCTTCTTAGGACCTCTAGACGCTATAATATT 240
6900 CCTCAGCAATGATGCCCTGGACTCTCCCTTCTTAGGACCTCTAGACGCTATAATATT 6959
QY 241 TTTTACTCTCTTTGGACCCCTGATCTTCAACTCTTCTTGTAGTTTCTCTCCAGAAT 300
6960 GCTACTCTCTTTGGACCCCTGATCTTTRACCTCTTGTAACTTTGTCTCTCTCCAGAAT 7019
Db

QY	301	TGAAGCTGTAAGAGCTACAAATAGTCTTCAAATGGAAACCCAGATGCAGTCCATGACTAA	360
Db	7020	CGAAGCTGTGTRAACTA-----CAAATGGAGCCCAAGATGCAGTCCAAAGACTAA	7067
QY	361	AATCTACCGTGGACCCCTGGACCGGCTGTCTAGACATATGCTCTGATGTTTAATGACATTTGA	420
Db	7068	GATCTACCGCAGACCCCTGGACCGGCTGVTAGCCACGATCTGATGTTTAATGACATCAA	7127
QY	421	AGTCACCCCTCCCGAGGAATCTCAATGTCACAAACCCCTACTACACTCCCAATTCAGTACG	480
Db	7128	AGGCACCCCTCCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG	7187
QY	481	AAGCAGTTAGAGCAGTTGTCTAGCCCAACCTCCCAACAGTACTTTGGGTTTTTCTGTGTTGAGA	540
Db	7188	AAGCAGTTAGAGCGGTSCTGGCCNACTCCCAACAGCAGTACTTTCTGTGTTGAGA	7247
QY	541	GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCNAGCCT	600
Db	7248	TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCYTAAGCCT	7307
QY	601	ANCTGGGAAGGTGACCCGATCCATCTTTAAACATGGGGCTTGCACACTTAGCTCACACCCG	660
Db	7308	AGSTGGGAAGGTGACCAATCCATCTTTAAACAGGGGCTTGCACACTTAGCTCACACCTG	7367
QY	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC	720
Db	7368	ACCAATCAGAGAGCTCACTAAATGCTAATGAGCAAGAGGAGTAAAGAAATAGCC	7427
QY	721	AATCATCTATTGCTGTGAGAGACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780
Db	7428	AATCATCTATTGCTGTGAGAGACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	7487
QY	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTGATATCGGAGCTCTGTTTT	840
Db	7488	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTGATATCGGAGCTCTGTTTT	7547
QY	841	CACCTCTATTTCACCTCTATTAAATCATGCAACTGCA	875
Db	7548	CATGCTATTTCACCTCTATTAAATCTTGCACTGCR	7582
RESULT 22			
ID	AAAS9215 standard; DNA; 7582 BP.		
XX	AAAS9215;		
AC	AAAS9215;		
DT	07-NOV-2000 (first entry)		
XX	Human endogenous retrovirus W (HERV-W) sequence.		
DE	Autoimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;		
KW	gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.		
XX	Human endogenous retrovirus.		
OS	Human endogenous retrovirus.		
XX	Key		
FH	Location/Qualifiers		
FT	LTR	1..120	
FT		/tag= a	
FT		/note= "R of 5' LTR"	
FT	LTR	121..575	
FT		/tag= b	
FT		/note= "U5 of 5' LTR"	
FT	primer_bind	579..596	
FT		/tag= c	
FT	CDS	581..7194	
FT		/tag= d	
FT		/note= "ORF1 env538"	
FT	CDS	7039..7194	
FT		/tag= e	
FT		/note= "ORF2 52 AA"	
FT	CDS	7112..7255	

[illegible]

||||| 7068 GATCTACCGACGACCCCTCGACCGCGCTGTAGCCACCATCTGATTTAATGACATCAA 7127
QY 421 AGTACCCCTCCCGAGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG 480
Db 7128 AGGCACCCCTCTCGAGAAATCTCAGCTGCACAAACCTCTACTACGCCCCCAATTCAGCAGG 7187
QY 481 AAGCAGTTAGACAGTTGTTCAGCCACCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
Db 7188 AAGCAGTTAGACGGTGTGCGGCAACCTCCCAACAGCAGCTTAGTGTTCCTGTTGAGA 7247
QY 541 GGGTGGACTCAGACAGACGACTAGCTGTGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 7248 TGGGGACTCAGACAGACGACTAGCTGTGATTTCTTAGGCTGATTAAGAAATCCVTAAGCCT 7307
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGSGGGCTTGCACCTTAGCTCAGACCCG 660
Db 7308 AGSTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGTTCACACCTG 7367
QY 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGGTAAAGCAATAGCC 720
Db 7368 ACCAATCAGAGAGCTCACTAAATGCTTAATAGGCAAAACAGAGGTAAAGCAATAGCC 7427
QY 721 AATCATCTATTGCTCAGACGACAGCGGGAAGGACAAAGGATTTGGGATATAAACTCAGGCA 780
Db 7428 AATCATCTATTGCTCAGACGACAGCGGGAAGGACAAATGATCGGATATAAAACCAAGTV 7487
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 840
Db 7488 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 7547
QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGCA 875
Db 7548 CATGCTATTTCACCTCTATTAAATCTTGCACTGCR 7582

RESULT 23

AA25660
ID AAX25660 standard; cDNA to mRNA; 1136 BP.
XX
AC AAX25660;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.C4C5.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
OS Human endogenous retrovirus.
XX
PN WO9902696-A1.
XX
PD 21-JAN-1999.
XX
PF 06-JUL-1998; 98WO-FR001442.
XX
PR 07-JUL-1997; 97FR-00008915.
XX
PA (INMR) BIO MERIEUX.
XX
PI Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;
XX
DR WPI; 1999-120897/10.
XX
PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.
XX
PS Claim 1; Page 59-60; 106pp; French.
XX
CC This sequence represents clone cl.C4C5 of the human endogenous retrovirus.

CC (HERV) W genome. The nucleic acids, their fragments or peptides encoded
CC by them are markers of autoimmune disease (e.g. multiple sclerosis,
CC rheumatoid polyarthritis, disseminated lupus erythematosus, insulin-
CC dependent diabetes and related pathologies) and of abnormal or
CC unsuccessful pregnancy and can be used as chromosomal markers for
CC susceptibility to these conditions, or proximity markers of genes
CC associated with this susceptibility

XX SQ Sequence 1136 BP; 336 A; 289 C; 241 G; 270 T; 0 U; 0 Other;

Query Match 54.3%; Score 721.2; DB 2; Length 1136;
Best Local Similarity 90.0%; Pred. No. 9.9e-226;
Matches 787; Conservative 0; Mismatches 75; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTAAACCCCAAAAGAGGGGGAACCTGTTTATTTT 60
Db 254 TCAAAATCGAAGAGCTTTAGACTTGTAAACCCCAAAAGAGGGGGAACCTGTTTATTTT 313
QY 61 AGGGGAAGAAATCTGTAGTATGTTAAATCAATCTCGAATCATTAATGAGAAATTAAGA 120
Db 314 AGGGGAAGAAATCTGTATTATGTTAAATCAATCCGGAATCGTCACTGAGAAATTAAGA 373
QY 121 AATTTGAGATCGAATATAATGTTAGACGACGAGACCTTCAAAACACTGCACCCCTGGGCGCT 180
Db 374 AATTCGAGATCGAATACACGTTAGACGAGAGCTTCGAAACACTGGACCCCTGGGCGCT 433
QY 181 CCTCAGCCAAATCGATCGCCCTGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATAT 240
Db 434 CTTACGCCAATGATGCGCCCTGATTTCTCCCTCTTTAGGACCTCTAGCAGCTATAATAT 493
QY 241 TTTACTCTCTTTTGGACCCCTGATTTCAACTTCTCTTTGTTAAAGTTTGTCTCTTCCAGAA 300
Db 494 GCTACTCTCTTTTGGACCCCTGATTTTAACTCTCTTTGTTAACTTGTCTCTTCCAGAA 553
QY 301 TGAAGCTGTAAGACTACAAATAGTTCTTCAATGGAACCCAGATCCAGTCCATGACTAA 360
Db 554 CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATCAGTCCAAAGACTAA 601
QY 361 AATCTACCGTGGACCCCTCGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
Db 602 GATCTACCCAGACCCCTCGACCGGCTGCTAGCCACCATCTGATGTTAATGACATCAA 661
QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCACAAACCCCTACTACCTTACGAGTTCAGTAGG 480
Db 662 AGGCACCCCTCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG 721
QY 481 AAGCAGTTAGACAGTTGTTCAGGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db 722 AAGCAGTTAGACGGGTGTCGGCCAACTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGA 781
QY 541 GGTGGACTCAGACAGGACTAGCTGGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 782 TGGGGACTCAGACAGGACTAGCTGGGATTTCTTAGGCTGACTAAGAAATCCCTAGCCT 841
QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGSGGGCTTGCACCTAGCTCAGACCCG 660
Db 842 AGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTGCAACTTAGTTTACACCTG 901
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGAGGTAAAGCAATAGCC 720
Db 902 ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGAGGTAAAGCAATAGCC 961
QY 721 AATCATCTATTGCTCAGACGACGAGGATTTGGGATATAAACTCAGGCA 780
Db 962 AATCATCTATTGCTCAGACGACGAGGAGCAATGATCGGATATAAAACCAAGTC 1021
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGAGCTCTGTTTT 840
Db 1022 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTT 1081
QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGC 874
Db 1082 CATGCTATTTCACCTCTATTAAATCTTGCACTGTC 1115

Db	494	GCTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCTTGTAACTTTGTCTCTTCCAGAAAT	553
Qy	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGNACCCGAGATGACGTCCATGACTAA	360
Db	554	CGAAGCTGTAAAGCTA-----CAAAATGGAGCCCAAGATGCAGTCCAAGACTAA	601
Qy	361	AATCTACCGTGGACCCCTGCTGACCGGCTGTAGACTATGCTCTGTATGTTAATGACATTTGA	420
Db	602	GATCTACCGAGACCCCTGAGACCGGCTGTAGCCACGATCTGATGTTAATGACATCA	661
Qy	421	AGTCACCCCTCCCGAGGAAATCTCAATGCAACACCCCTACTACACTCCAAATTCAGTAGG	480
Db	662	AGCACCCCTCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG	721
Qy	481	AAGCAGTTAGAGCAGTTGTCAGCCAACTCCCAACAGTACTTTGGGTTTCTCTGTTGAGA	540
Db	722	AAGCAGTTAGAGCGGTGCTGGCCAACTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGA	781
Qy	541	GGGTGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	600
Db	782	TGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT	841
Qy	601	ANCTGGGAAGTGACCGCATCTCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660
Db	842	AGCTGGGAAGTGACCCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG	901
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC	720
Db	902	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC	961
Qy	721	AATCATCTATTGCTGTAGAGCAGCGGGGAAGCAAGGATTTGGGATATATAACTCAGGCA	780
Db	962	AATCATCTATTGCTGTAGAGCAGCGAGGAGGACAAATGATCGGGATATAAACCCCAAGTC	1021
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGCTCCCTCCCATTTGATGAGGAGCTCTGTTT	840
Db	1022	TTGAGCGCGCAACGGCAACCCCTTTGGGTCCTCTCTCTCTTTGATGGGAGCTCTGTTT	1081
Qy	841	CACCTCTATTCACTCTATTAAATCATGCAACTGC	874
Db	1082	CATGCTATTCACTCTATTAAATCTTCAGCTGC	1115
XX	RESULT 25		
XX	AAX25661		
XX	ID AAX25661	standard; cDNA to mRNA; 2782 BP.	
XX	AC AAX25661;		
XX	XX		
XX	21-MAY-1999	(first entry)	
XX	XX	Human endogenous retrovirus W clone cl.PH74.	
XX	XX	Clone; human endogenous retrovirus; genome; autoimmune disease;	
XX	XX	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;	
XX	XX	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.	
XX	OS	Human endogenous retrovirus.	
XX	XX	WO9902696-A1.	
XX	XX		
XX	PD	21-JAN-1999.	
XX	XX		
XX	PF	06-JUL-1998; 98WO-FR001442.	
XX	XX		
XX	PR	07-JUL-1997; 97FR-00008815.	
XX	XX	(INMR) BIO MERIEUX.	
XX	PA	Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;	
XX	PI	WPI; 1999-120897/10.	
XX	XX		
XX	DR		
XX	XX		

Db	494	GCTACTCCTCTTTGGACCCCTGTATCTTTAAACCTCTTGTAACTTTGTCTCTTCCAGAAAT	553
Qy	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAATGGNACCCGAGATGACGTCCATGACTAA	360
Db	554	CGAAGCTGTAAAGCTA-----CAAAATGGAGCCCAAGATGCAGTCCAAGACTAA	601
Qy	361	AATCTACCGTGGACCCCTGCTGACCGGCTGTAGACTATGCTCTGTATGTTAATGACATTTGA	420
Db	602	GATCTACCGAGACCCCTGAGACCGGCTGTAGCCACGATCTGATGTTAATGACATCA	661
Qy	421	AGTCACCCCTCCCGAGGAAATCTCAATGCAACACCCCTACTACACTCCAAATTCAGTAGG	480
Db	662	AGCACCCCTCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG	721
Qy	481	AAGCAGTTAGAGCAGTTGTCAGCCAACTCCCAACAGTACTTTGGGTTTCTCTGTTGAGA	540
Db	722	AAGCAGTTAGAGCGGTGCTGGCCAACTCCCAACAGCAGCTTAGGTTTCTCTGTTGAGA	781
Qy	541	GGGTGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT	600
Db	782	TGGGGAGCTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT	841
Qy	601	ANCTGGGAAGTGACCGCATCTCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660
Db	842	AGCTGGGAAGTGACCCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG	901
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC	720
Db	902	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC	961
Qy	721	AATCATCTATTGCTGTAGAGCAGCGGGGAAGCAAGGATTTGGGATATATAACTCAGGCA	780
Db	962	AATCATCTATTGCTGTAGAGCAGCGAGGAGGACAAATGATCGGGATATAAACCCCAAGTC	1021
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGCTCCCTCCCATTTGATGAGGAGCTCTGTTT	840
Db	1022	TTGAGCGCGCAACGGCAACCCCTTTGGGTCCTCTCTCTCTTTGATGGGAGCTCTGTTT	1081
Qy	841	CACCTCTATTCACTCTATTAAATCATGCAACTGC	874
Db	1082	CATGCTATTCACTCTATTAAATCTTCAGCTGC	1115
XX	RESULT 25		
XX	AAX25661		
XX	ID AAX25661	standard; cDNA to mRNA; 2782 BP.	
XX	AC AAX25661;		
XX	XX		
XX	21-MAY-1999	(first entry)	
XX	XX	Human endogenous retrovirus W clone cl.PH74.	
XX	XX	Clone; human endogenous retrovirus; genome; autoimmune disease;	
XX	XX	multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;	
XX	XX	disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.	
XX	OS	Human endogenous retrovirus.	
XX	XX	WO9902696-A1.	
XX	XX		
XX	PD	21-JAN-1999.	
XX	XX		
XX	PF	06-JUL-1998; 98WO-FR001442.	
XX	XX		
XX	PR	07-JUL-1997; 97FR-00008815.	
XX	XX	(INMR) BIO MERIEUX.	
XX	PA	Beseme F, Blond J, Bouton O, Mandrand B, Mallet F;	
XX	PI	WPI; 1999-120897/10.	
XX	XX		
XX	DR		
XX	XX		

PT New nucleic acid sequences from human endogenous retrovirus-W - expressed
PT exclusively in placenta and useful in diagnosis and therapy of autoimmune
PT disease, and abnormal or failed pregnancy.

XX Claim 1; Page 60-63; 106pp; French.

XX This sequence represents clone cl.PH74 of the human endogenous retrovirus
XX (HERV) W genome. The nucleic acids, their fragments or peptides encoded
XX by them are markers of autoimmune disease (e.g. multiple sclerosis,
XX rheumatoid polyarthritits, disseminated lupus erythematosus, insulin-
XX dependent diabetes and related pathologies) and of abnormal or
XX unsuccessful pregnancy and can be used as chromosomal markers for
XX susceptibility to these conditions, or proximity markers of genes
XX associated with this susceptibility

XX Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 53.6%; Score 712.6; DB 2; Length 2782;
Best Local Similarity 89.4%; Pred. No. 1.1e-222;
Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGAGGGGAACTGTTATTTT 60
DB 1902 TCGAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGAGGGGAACTGTTATTTT 1961
QY 61 AGGGGAAGATCTGTAGTATGTTAATCAATCTGGAATCACTAGCAAGATTAAGA 120
DB 1962 AGGGGAAGATCTGTATGTTAATCAATCGGAATCGTCACTGAGAAAGTTGAAGA 2021
QY 121 AATTGAGATCGAATATAATGTAGCAGAGGACCTTCAAAACACTGCACCTGGGGCCT 180
DB 2022 AATTGAGATCGAATATAATGTAGCAGAGGACCTTCAAAACACTGCACCTGGGGCCT 2081
QY 181 CCTCAGCCAAATGAGTGCCTGAGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
DB 2082 CCTCAGCCAAATGAGTGCCTGAGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 2141
QY 241 TTTACTCTCTTTGGACCGTATCTCACTTCTTGAAGTTGCTCTCTCCAGAAAT 300
DB 2142 GCTACTCTCTTTGGACCGTATCTCACTTCTTGAAGTTGCTCTCTCCAGAAAT 2201
QY 301 TGAAGCTGTAAAGCTCAAAATAGTTCTTCAAATGGAACCCAGATGCAGTCCATGACTAA 360
DB 2202 CGAAGCTGTAAAGCTCAAAATAGTTCTTCAAATGGAACCCAGATGCAGTCCATGACTAA 2249
QY 361 AATCTACCGTGGACCCCTGAGCCGCTGTAGACTATGCTGTGATGTTAATGACATTTGA 420
DB 2250 GATCTACCGCAGACCCCTGAGCCGCTGTAGCCACGATCTGATGTTAATGACATCAA 2309
QY 421 AGTCACCCCTCCGAGGAATCTCACTGCACACCCCTACTACCTCCCAATTCAGTAGG 480
DB 2310 AGGCACCCCTCCGAGGAATCTCACTGCACACCCCTACTACCTCCCAATTCAGTAGG 2369
QY 481 AAGCAGTTAGCAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTTCCTGTTGAGA 540
DB 2370 AAGCAGTTAGCAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTTCCTGTTGAGA 2429
QY 541 GGTGTGACTGAGAGACAGGACTAGTGGATTTCTAGGCTGACTAAGATCCCNAGCCT 600
DB 2430 TGGGGGACTGAGAGACAGGACTAGTGGATTTCTAGGCTGACTAAGATCCCNAGCCT 2489
QY 601 ANCTGGGAAGGTGACCCGATCTTTAAACATGAGGCTTGAACCTAGCTCAGACCCG 660
DB 2490 AGTGGGAAGGTGACCCGATCTTTAAACATGAGGCTTGAACCTAGCTCAGACCCG 2549
QY 661 ACCAATCAGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 720
DB 2550 ACCAATCAGAGAGCTCACTAAATATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 2609
QY 721 AATCATCTATTGCTGTAGAGACAGCGGGAGGCAAGAGATTTGGATATATAAATCAGGCA 780
DB 2610 AATCATCTATTGCTGTAGAGACAGCGGGAGGCAATGATCGGGATATATAAATCAGGAT 2669

QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTT 840
DB 2670 TTGAGCGGCAACGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTT 2729
QY 841 CACTCTATTTCACTCTATTAATCATGCAACTGCA 875
DB 2730 CATGCTATTTCACTCTATTAATCATGCAACTGCA 2764

RESULT 26

AAAS9211
ID AAAS9211 standard; DNA; 2782 BP.
AC AAAS9211;
XX
XX 07-NOV-2000 (first entry)
XX
XX 5' non coding, 3' pol gene, and 3' non coding sequences of HERV-W.
XX Autolimmune disease; retrovirus; human endogenous retrovirus W; HERV-W;
XX gag gene; pregnancy; multiple sclerosis; T cell proliferation; ss.
XX Homo sapiens.
XX WO200043521-A2.
XX
XX 27-JUL-2000.
XX
XX 21-JAN-2000; 2000WO-FR000144.
XX
XX 21-JAN-1999; 99FR-00000888.
XX (INMR) BIO MERIEUX.
XX Paranhos-Baccala G, Mallet P, Voisset C;
XX WPI; 2000-499229/44.
XX
XX New nucleic acid from human endogenous retrovirus, useful e.g. for
XX diagnosis of autoimmune disease and complications of pregnancy, contains
XX at least part of the gag gene.
XX Disclosure; Page 46-47; 53pp; French.

XX The present sequence represents an endogenetic retroviral nucleic acid
XX fragment, which is associated with an autoimmune disease, and is
XX integrated into the human genome. The fragment is originally derived from
XX a novel retrovirus, human endogenous retrovirus W (HERV-W). The HERV-W
XX retrovirus is associated with autoimmune disease, failure of pregnancy or
XX disorders of pregnancy. The nucleic acid fragment, or proteins derived
XX from it, are useful for diagnosis of autoimmune disease (specifically
XX multiple sclerosis) and for monitoring pregnancy. The nucleic acid
XX fragments may also be used for in situ labelling of isolated chromosomes,
XX while the transcription product can be used to study or monitor T cell
XX proliferation in vitro
XX Sequence 2782 BP; 741 A; 767 C; 565 G; 709 T; 0 U; 0 Other;

Query Match 53.6%; Score 712.6; DB 3; Length 2782;
Best Local Similarity 89.4%; Pred. No. 1.1e-222;
Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTATTTT 60
DB 1902 TCGAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAGAGAGGGGAACTGTTATTTT 1961
QY 61 AGGGGAAGATCTGTTAGTATGTTAATCAATCTGGAATCATTAAGTAAAGA 120
DB 1962 AGGGGAAGATCTGTTATGTTAATCAATCCGAATCGTCACTGAGAAAGTTGAAGA 2021
QY 121 AATTGAGATCGAATATAATGTAGCAGAGGACCTTCAAAACACTGCACCTGGGGCCT 180
DB 2022 AATTGAGATCGAATATAATGTAGCAGAGGACCTTCAAAACACTGCACCTGGGGCCT 2081

```
QY 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTAATAATT 240
Db 2082 CCTCAGCGATGGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTAATAATT 2141
QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTGTTAAAGTTTGTCTCTCCAGAA 300
Db 2142 GCTACTCTCTTTGGACCTGTATCTTGGACCTCTTGTAACTTGTCTCTCCAGAA 2201
QY 301 TGAAGCTGAAGCTTACAATAAGTCTTCAATAGGAACCCAGATGAGTCCATGACTAA 360
Db 2202 CGAAGCTGTGAACATA-----CAAATGGAGCCCAAGATGAGTCCCAAGACTAA 2249
QY 361 AATCTACCGTGGACCCCTGGACCGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2250 GATCTACCGCAGACCCCTGGACCGCTGTAGCCACGATCTGATGTTAATGACATCAA 2309
QY 421 AGTCACCCCTCCGAGGAATCTCACTGCACACCCCTACTACACTCCAATTCAGTAGG 480
Db 2310 AGGCACCCCTCTGAGGAATCTCAGCTGCACAACTCTACTAGCCCAATTCAGCAGG 2369
QY 481 AAGCAGTTTAGACAGCTGTCTAGCCAACTCTCCCAACAGTACTTGGTCTTCTGTTGAGA 540
Db 2370 AAGCAGTTTAGACGCTGTCTAGCCAACTCTCCCAACAGTACTTGGTCTTCTGTTGAGA 2429
QY 541 GGGTGGACTGAGACAGACGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
Db 2430 TGGGGGACTGAGACAGACGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCTTAAGCCT 2489
QY 601 ANCTGGGAGGTGACCGCATCATCTTTAAACATGGGCTTGCACTTAGCTCACACCGG 660
Db 2490 AGGTGGGAGGTGACCATCATCTTTAAACATGGGCTTGCACTTAGCTCACACCTG 2549
QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 2550 ACCAATCAGAGAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCC 2609
QY 721 AATCATCTATTCCTGAGAGCAGCGGAGGACAAGATTTGGATATTAACCTCAGGCA 780
Db 2610 AATCATTTATTGCTGAGAGCAGCAGGAGGACAATGATCGGATATAAACCCCAAGTT 2669
QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTGTTT 840
Db 2670 TTCGAGCCGCAACGCAACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGTGTTT 2729
QY 841 CACTCTATTTCATCTATTAAATCATGCAACTGCA 875
Db 2730 CATGCTATTTCATCTATTAAATCTTGCAACTGCA 2764

RESULT 27
AAH20069
ID AAH20069 standard; DNA; 2782 BP.
AC AAH20069;
XX
DT 08-AUG-2001 (first entry)
XX
DE HERV-W envelope protein G encoding nucleic acid.
XX
KW Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW envelope protein; multiple sclerosis-related superantigen; vaccine;
KW surface antigen; transmembrane; multiple sclerosis; neuroprotective;
KW antisense-therapy; autoimmune disorder; ds.
XX
OS Human endogenous retrovirus.
XX
PH Key Location/Qualifiers
FT 5'UTR 1..762 a
FT 763..2379
FT /*tag= b
FT /*product= "HERV-W envelope protein G"
FT
```

```
FT /transl_except= (pos:790..792,aa:Phe)
FT /transl_except= (pos:793..795,aa:Thr)
FT /transl_except= (pos:812..814,aa:Leu)
FT /transl_except= (pos:818..820,aa:Ser)
FT /transl_except= (pos:862..864,aa:Tyr)
FT /transl_except= (pos:865..867,aa:Gln)
FT /transl_except= (pos:1174..1176,aa:Arg)
FT /transl_except= (pos:1441..1443,aa:Leu)
FT /transl_except= (pos:1903..1905,aa:Gln)
FT /transl_except= (pos:2017..2019,aa:Lys)
FT /transl_except= (pos:2026..2028,aa:Arg)
FT /transl_except= (pos:2044..2046,aa:Arg)
FT /transl_except= (pos:2089..2091,aa:Gln)
FT /transl_except= (pos:2170..2172,aa:Asn)
FT 2380..2782
FT /*tag= c
XX
XX MO200131021-A1.
XX
XX 03-MAY-2001.
XX
XX 30-OCT-2000; 2000WO-EF010659.
XX
XX 28-OCT-1999; 99EP-00402690.
XX
XX (UYGE-) UNIV GENEVE.
XX
XX Conrad B, Mach B;
XX
XX WPI; 2001-316336/33.
XX P-PSDB; AAB75138.
XX
XX New human retrovirus HERV-W ENV proteins/peptides having superantigen
XX activity useful for diagnosing and treating multiple sclerosis.
XX
XX Claim 13; Fig 9; 94pp; English.
XX
XX On the basis of the PBS t-RNA motif used for the classification of human
XX endogenous retrovirus (HERVs) the full length endogenous provirus which
XX was located on the long arm of human chromosome 7 (7q21-22) has been
XX designated HERV-W. The present invention describes proteins or peptides
XX (I) having superantigen (SAG) activity comprising the ENV protein (ENV)
XX of HERV-W, the surface protein (SU) and can be used in: vaccines; antisense-
XX therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are
XX useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
XX disorders. (I) are also useful for identifying superantigen associated
XX recovering), capable of binding to a retroviral superantigen and optionally
XX with MS, substances capable of blocking SAG activity and substances
XX capable of blocking transcription or translation of HERV-W retroviral
XX superantigen. A protein or peptide derived from (I), modified to be
XX devoid of SAG activity and being capable of generating an immune response
XX against HERV-W retroviral SAG is useful in therapy. Nucleic acid
XX molecules encoding (I) are useful as vaccines against MS. Substances
XX capable of blocking SAG activity, capable of binding to a retroviral
XX superantigen associated with MS, or capable of blocking transcription or
XX translation of HERV-W retroviral superantigen for use in treating or
XX preventing MS, obtained using (I) are useful for the treatment and
XX prevention of MS. (I) and nucleic acids encoding them are useful for
XX diagnosing autoimmune disease. The present sequence encodes the
XX specifically claimed envelope protein of HERV-W designated G
XX
XX Sequence 2782 BP; 741 A; 768 C; 564 G; 709 T; 0 U; 0 Other;
XX
XX Query Match 53.6%; Score 712.6; DB 5; Length 2782;
XX Best Local Similarity 89.4%; Pred. No. 1.1e-222;
XX Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;
XX
XX QY 1 TCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAGAGGGGAACTGTTTATTTT 60
XX Db 1902 TCGAAATCGAAGAGCTTTAGACTTCTTAACCGCTGAGAGGGGAACTGTTTATTTT 1961
XX QY 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCAATTAAGGAAATGTTAAAGA 120
```

Db 1962 AGGGAAGAAATGCTGTTATATGTTATCAATCCGGAATCGTCACTGAGAAATGTAAGA 2021
Qy 121 AATTTGAGATCGAATATATATAGACAGAGACCTTCAAAACACTGCAACCTGGGGCT 180
Db 2022 AATTCAGATCGAATACAAACGATATAGCAGAGAGCTTCGAAACACTGGACCTGGGGCT 2081
Qy 181 CCTCAGCAATGATGCGCTGCACTCTCCCTCTTAGACCTCTAGCAGCTATAATATT 240
Db 2082 CCTCAGCCGATGATGCGCTGGAATCTCCCTCTTAGACCTCTAGCAGCTATAATATT 2141
Qy 241 TTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGAAGTTTGTCTCTTCCAGAT 300
Db 2142 GCTACTCTCTTTGGACCTGTATCTTTGACCTCTCTTGAAGTTTGTCTCTTCCAGAT 2201
Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAATGACTAA 360
Db 2202 CGAAGCTGTAACTA-----CAAAATGGAGCCCAAGATGCAATGCAAGACTAA 2249
Qy 361 AATCTACCTGGAACCTTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2250 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACCATCTGATGTTAATGACATCA 2309
Qy 421 AGTCACCTCCCGAGGAAATCTCAACTGCACAAACCTTACTACACTTCAATTCAGTAGG 480
Db 2310 AGGACCCCTCTTGAGGAAATCTCAGCTGCACAACTTACTAGCCCAATTCAGCAGG 2369
Qy 481 AAGCAGTTAGCAGATTGTGAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db 2370 AAGCAGTTAGCAGGTTGGTGGGCAACCTCCCAACAGCAGTACTTGGGTTTCTCTGTTGAGA 2429
Qy 541 GGTGTAGCTGAGAGACAGCTAGCTGGATTCTTCTAGGCTGACTAGAAATCCCNAGCCT 600
Db 2430 TGGGGGACTGAGAGACAGGACTAGCTGGATTCTTCTAGGCTGACTAGAAATCCCTTAAAGCCT 2489
Qy 601 ANCTGGGAAGGTACCGCATCTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660
Db 2490 AGTGGGAAGGTGACACATCTCACTTAAACAGGGGCTTGCAACTTAGCTCACACCTG 2549
Qy 661 ACCAATCAGAGAGCTCACTAAAATGCTAATCAGGCAAAACAGAGAGTAAAGCAATAGCC 720
Db 2550 ACCAATCAGAGAGCTCACTAAAATGCTAATCAGGCAAAACAGAGAGTAAAGCAATAGCC 2609
Qy 721 AATCATCTATTGCTGAGACACAGCGGGAAGGACAGAGATTGGGATATAAATCAGGCA 780
Db 2610 AATCATTTATTGCTGAGACACAGCAGGAGGAGCAATGATCGGGATATAAATCAGGCA 2669
Qy 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATTTGATGGGAGCTCTGTTTT 840
Db 2670 TTGAGCGGCAACGGCAACCCCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTT 2729
Qy 841 CACTCTATTTCTCTATTAAATCATGCAACTGCA 875
Db 2730 CATGCTATTTCTCTATTAAATCTTGCAACTGCA 2764

RESULT 28

AA114608

ID AA114608 standard; DNA; 1894 BP.

XX AC AA114608;

XX AC AA114608;

XX 12-OCT-2001 (first entry)

XX Probe #4541 for gene expression analysis in human cervical cell sample.

DE DE Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer; ss.

XX KW Homo sapiens.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX XX

PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000670.
PF 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
PT Claim 25; SEQ ID NO 4541; 487bp; English.
PS The present invention relates to human single exon nucleic acid probes
XX (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
SQ Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 120
Db 961 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 1020
Qy 61 AGGGGAAGAAATGCTGTTATGTTTAAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 120
Db 1021 AGAGGAAATGCTGTTATGTTTAAATCAATCTGGAATCATTAATCTGAGAAAGTAAAGA 1080
Qy 121 AATTTGAGATCGAATATAATAGTACAGAGGAGCTTCAAAACACTGCAACCTGGGGCCT 180
Db 1081 AATTTGAGATCGAATATAATAGTACAGAGGAGCTTCAAAACACTGCAACCTGGGGCCT 1140
Qy 181 CCTCAGCAATGATGCGCTGCACTCTCCCTCTTTAGACCTCTAGCAGCTATAATATT 240
Db 1141 CCTCAGCAATGATGCGCTGCACTCTCCCTCTTTAGACCTCTAGCAGCTATAATATT 1200
Qy 241 TTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGAAGTTTGTCTCTTCCAGAT 300
Db 1201 GTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGAAGTTTGTCTCTTCCAGAT 1260
Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCAGATGCAATGACTAA 360
Db 1261 CGAAGCAGTAAACTACAAATCGTTCTTCAATGGAGCCCGAGATGCAATGACTAA 1320
Qy 361 AATCTACCTGGAACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 1321 AATCTACCGGACCCCTGGACCGGCTGCTAGCCCATGCTCTGATGTTAATGACATCAA 1380
Qy 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCACAAACCTTACTACCTCCCAATTCAGTAGG 480
Db 1381 AGGACCCCTCCCGAGGAAATCTCAACTGCACAAACCTTACTACCTCCCAATTCAGTAGG 1440
Qy 481 AAGCAGTTAGCAGATTGTGAGCCAAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540

CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTTGCTTAACCGCAAAAGAGGGGACCTGTTTATTTT 60
Db 961 TCAAAATCGAAGAGCTTTAGACTTTGCTTAACCGCAAAAGAGGGGACCTGTTTATTTT 1020
Qy 61 AGGGGAAGATGCTGTTAGTATCTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
Db 1021 AGAGGAAAATGCTGTTGTTATGTTTAAATCAATCCGGAATCATCCCGAAGAGTTAAAGA 1080
Qy 121 AATTTGAGATCGAATATAATGTAGACAGAGGACCTTCAAAAACATGACACCTGGGCGCT 180
Db 1081 AATTCAGGTGGAATATAACGTAGAGCAAGAGAGCTGCAAAAACATGGAACCTGGGCGCT 1140
Qy 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
Db 1141 CCTCAGCAATGGATGCCCTGGATTTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 1200
Qy 241 TTTTACTCTCTTTGGACCTGTATCTTCAACTCTCTTTGTTAAGTTTCTCTTCCAGAAT 300
Db 1201 GTTACTCTCTTTGGACCTGTATCTTAACTCTCTTTGTTAAGTTTCTCTTCCAGAAT 1260
Qy 301 TGAAGCTGTAAGCTACAAATAGTTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
Db 1261 CGAAGCAGTAAACTCAAAATCGTTCTTCAAAATGGAGCCCGAGATGCGATGAGTAA 1320
Qy 361 AATCTACGTTGGACCCCTGGACCGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
Db 1321 AATCTACGCGGACCCCTGGACCGCTGCTAGCCCATGCTCTGATGTTAATGACATCAA 1380
Qy 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCAACACCCCTACTACCTCCAAATTCAGTAGG 480
Db 1381 AGGCACCCCTCCCGAGGAAATCTCAACTGCAACACCTCTACTACGCCCCCAATTCAGCAGG 1440
Qy 481 AAGCAGTTAGACAGAGTTGTCAGCCAACTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db 1441 AAGCAGTTAGAGTGGTTGTTGGCCAACTCCCAACAGCAGTGGGTTTCTCTGTTGAGA 1500
Qy 541 GGGTGGCTGAGACAGGAGTACTGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAGCCT 600
Db 1501 GGGGGGCTGAGACAGGAGTAACTAGATTTCTTAGACCACTAAGAAATCCCTAAGACT 1560
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCAACTTTAGCTGACACCG 660
Db 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCAACTTTAGCTGACGCCA 1620
Qy 661 ACCAATC-----AGAGAGCTCACTAAATGCTATCAGGCAAAACAGGAGGTAAA 711
Db 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTAATTTAGGCAAAACAGGAGATAA 1680
Qy 712 GCAATAGCCAATCATCTATTGCTGAGACAGCAGCGGGAAGGACAAGGATTTGGGATATAA 771
Db 1681 GAAATAGCCAATCATCTGTTGCTTGACAGCAGCAGGAGGAGCAATGATCGGATATAA 1740
Qy 772 ACTCAGGCATTTCAAGCCAGCAACAGACCCCTTTGGGTCCCTCCCATTTGTATGGGAG 831
Db 1741 ACCCAGGCATTCGAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGTATGGGAG 1800
Qy 832 CTCTGTTTCACTCTATT 849
Db 1801 CTCTGTTTCACTCTATT 1818
RESULT 29
ID ABA56337 standard; DNA; 1894 BP.
XX ABA56337;
XX ABA56337;
DT 01-FEB-2002 (first entry)
XX Human foetal liver single exon nucleic acid probe #4642.
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human foetal liver.
XX Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed

Db 1441 AAGCAGTTAGAGTGGTTCTTGGCCAACTCCCAACAGCAGTGGGTTTCTCTTTGAGA 1500
Qy 541 GGGTGGACTGAGACAGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 600
Db 1501 GGGGGGACTGAGACAGGAAATAACTAGATTTCTTAGCACTAAGAAATCCCTAAGACT 1560
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCAACTTTAGCTGACACCG 660
Db 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCAACTTTAGCTGACGCCA 1620
Qy 661 ACCAATC-----AGAGAGCTCACTAAATGCTATCAGGCAAAACAGGAGGTAAA 711
Db 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTAATTTAGGCAAAACAGGAGATAA 1680
Qy 712 GCAATAGCCAATCATCTATTGCTGAGACAGCAGCGGGAAGGACAAGGATTTGGGATATAA 771
Db 1681 GAAATAGCCAATCATCTGTTGCTTGACAGCAGCAGGAGGACANTGATCGGATATAA 1740
Qy 772 ACTCAGGCATTTCAAGCCAGCAACAGACCCCTTTGGGTCCCTCCCATTTGTATGGGAG 831
Db 1741 ACCCAGGCATTCGAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGTATGGGAG 1800
Qy 832 CTCTGTTTCACTCTATT 849
Db 1801 CTCTGTTTCACTCTATT 1818
RESULT 29
ID ABA56337 standard; DNA; 1894 BP.
XX ABA56337;
XX ABA56337;
DT 01-FEB-2002 (first entry)
XX Human foetal liver single exon nucleic acid probe #4642.
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human foetal liver.
XX Claim 1; SEQ ID NO 4642; 639pp + Sequence Listing; English.
XX The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed

XX AAI35980;
 AC 17-OCT-2001 (first entry)
 DT Probe #4666 used to measure gene expression in human placenta sample.
 XX Probe #4666 used to measure gene expression in human placenta sample.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder; ss.
 KW Homo sapiens.
 XX WO200157272-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000663.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 XX Claim 25; SEQ ID NO 4666; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders
 XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
 SQ

Query Match 53.0%; Score 704.8; DB 4; Length 1894;
 Best Local Similarity 89.7%; Pred. No. 3.3e-220;
 Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;

QY 1 TCAAAATCGAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTTTATTTT 60
 DB 961 TCAAAATCGAGAGCTTTAGACTTGTCTAACCGCCAAAGAGGGGAACTTTATTTT 1020

QY 61 AGGGGAAGATGCTGTAGTATCTTAATCAATCTGGAATCATTACTGAGAAAGTTAAAGA 120
 DB 1021 AGAGGAAATGCTGTGTATGTATTAATCAATCCGAATCATCCGAGAAAGTTAAAGA 1080

QY 121 AATTTCAGATCGAATATAATGTAGACAGAGGACCTTCAAAAACACTGCACCCCTGGGCGCT 180
 DB 1081 AATTCAAGTCGGAATATAACGTAGACAAAGAGGCTGCAAAAACACTGCACCCCTGGGCGCT 1140

QY 181 CCTCAGCAATGATGCTGCTGGACTTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 240
 DB 1141 CCTCAGCAATGATGCTGCTGGACTTCCCTCTTTAGGACCTCTAGCAGCTATAATTT 1200

QY 241 TTTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAGTTGTCTCTTCCAGAAT 300
 DB 1201 GTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAGTTGTCTCTTCCAGAAT 1260

QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
 DB 1261 CGAAGCAGTAAACTACAAATCGTCTTCAAAATGGAACCCAGATGAGTCCATGAGTAA 1320

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
 DB 1321 AATCTACCGACGACCCCTGGACCGGCTGTAGCCATGCTCTGATGTTAATGACATCAA 1380

QY 421 AGTCACCCCTCCGAGGAAATCTCACTGACACACCCCTACTACACTCCAATTCAGTAGG 480
 DB 1381 AGGCACCCCTCCGAGGAAATCTCACTGACACACCCCTACTACACTCCAATTCAGTAGG 1440

QY 481 AAGCAGTTAGAGCAGTGTCTGACGCCAACCTCCCAACAGTACTTTGGGTTTCTCTGTTGAGA 540
 DB 1441 AAGCAGTTAGAGTGTGTGTCGCCAACCTCCCAACAGCAGTTGGGTTTCTCTGTTGAGA 1500

QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCCTTAGCTGACTAAGAAATCCNAAGCCT 600
 DB 1501 GGGGGGACTGAGAGACAGGAAATACTAGATTTCCTTAGACCAACTAAGAATCCCTAAGACT 1560

QY 601 ANCTGGGAAGTGACCGCATCCATCTTTAAACATGGGGCTTGCACACTTAGCTCACACCCG 660
 DB 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCACACTTAGCTCACGCCCA 1620

QY 661 ACCAATC-----AGAGAGCTCACTAAATATGCTAATCAGGCAAAAACAGGAGTAAA 711
 DB 1621 ACCAATCAGATACTAAAGAGAGCTCTAATAATGCTAATTAGGCAAAAACAGGAGATAAA 1680

QY 712 GCAATAGCCAATCATCTATTGCTTGAGAGCAGCGGGAAGGACAAGGATTGGGATATAA 771
 DB 1681 GAAATAGCCAATCATCTGTTGCTGACAGCAGCAGGAGGACAATGATCGGGATATAA 1740

QY 772 ACTCAGGCAATCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCTCCCTATGATGGGAG 831
 DB 1741 ACCCAGGCAATTCGAGCCAGCTACAGTACCTCTTTGGTCCCTCCCTCTTTGATGGGAG 1800

QY 832 CTCTGTTTCACTCTATT 849
 DB 1801 CTCTGTTTCACTCTATT 1818

RESULT 31
 ABM45822
 ID ABM45822 standard; DNA; 1894 BP.
 XX AC ABA45822;
 XX AC ABA45822;
 XX 01-FEB-2002 (first entry)
 XX Human breast cell single exon nucleic acid probe #4517.
 DE Human; microarray; single exon probe; gene expression; breast; disease;
 KW cancer; ss.
 XX Homo sapiens.
 XX WO200157271-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000662.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-496933/54.
 XX New spatially-addressable set of single exon nucleic acid probes, useful

PT for measuring gene expression in sample derived from human breast,
PI comprises number of single exon nucleic acid probes.
XX
PS Claim 1; SEQ ID NO 4517; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
XX
Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Best Local Similarity 89.7%; Pred. NO. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
QY 1 TCMAAATCGAAGAGCTTTAGACTTCTAATCGCGCAAGAGGGGAACTGTTATTTT 60
DB 961 TCMAAATCGAAGAGCTTTAGACTTCTAATCGCGCAAGAGGGGAACTGTTATTTT 1020
QY 61 AGGGAAGAATGCTTTAGTATGTTAATCAATCGGAATCAATCGGAAAGTAAAGA 120
DB 1021 AGAGGAAATGCTGTTGTTATGTTAATCAATCGGAAATCAATCGGAAAGTAAAGA 1080
QY 121 AATTGAGATCGAATATAATGATAGACAGAGGACCTTCAAAACATCGACCTGGGCGCT 180
DB 1081 AATTCAAGGTCGAATATAACGTAGACAGAGGAGCTGCAAAACATCGACCTGGGCGCT 1140
QY 181 CCTCAGCAATGAGTGGCTCGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
DB 1141 CCTCAGCAATGAGTGGCTCGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 1200
QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTGTTAAGTTTGTCTTTCAGAAAT 300
DB 1201 GTTACTCTCTTTGGACCTGATCTTCAACTTCTGTTAAGTTTGTCTTTCAGAAAT 1260
QY 301 TGAAGCTGTAAGCTTACAATAGTTCTTCAATGGAACCCGAGATGCAATGACTAA 360
DB 1261 CGAAGCAGTAAACTACAAATCGTTCTTCAATGGAACCCGAGATGCAATGACTAA 1320
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTGA 420
DB 1321 AATCTACCGGACCCCTGGACCGGCTGCTAGCCATGCTCTGATGTTAATGACTAA 1380
QY 421 AGTCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTACACTCAATTCAGTAGG 480
DB 1381 AGGCAACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTACACTCAATTCAGTAGG 1440
QY 481 AAGCAGTTAGACAGTCTGACGCAACCTCCCAACAGTACTTGGGTTTCTGTTTGA 540
DB 1441 AAGCAGTTAGAGTGGTTGTTGGCCAACTCCCAACAGTACTTGGGTTTCTGTTTGA 1500
QY 541 GGGTGGACTGAGACAGAGTCTGAGTGGATTTCTTAGGCTGACTAAGAAATCCCAAGGCT 600
DB 1501 GGGGGAAGTACAGACAGGAAATCAATAGTTTCTTAGACCACTAAGAAATCCCTAAGACT 1560
QY 601 ANCTGGGAAGGTACCGCATCTTTAAACATGGGGTGTGCAACTAGCTCACACCG 660
DB 1561 AGCTGGGAAGGTACCGCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTCAGGCCA 1620

QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGTAAA 711
DB 1621 ACCAATCAGATCTAAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGATAA 1680
QY 712 GCAATAGCCCAATCATCTATTGCTGAGACACAGCGGGAAGGACAGGATTGGGATATAA 771
DB 1681 GAAATAGCCCAATCATCTATTGCTGAGACACAGCGGGAAGGACATATCGGGATATAA 1740
QY 772 ACTCAGCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAG 831
DB 1741 ACCAGGCAATTCGAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGTATGGAG 1800
QY 832 CTCTGTTTTCACTCTATT 849
DB 1801 CTCTGTTTTCACTCTATT 1818
RESULT 32
ABA25978
ID ABA25978 standard; DNA; 1894 BP.
XX
AC ABA25978;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #4444 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human hearts.
XX
PS Claim 1; SEQ ID NO 4444; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
XX

Query Match	53.0%	Score 704.8	DB 4	Length 1894
Best Local Similarity	89.7%	Pred. No. 3.3e-220		
Matches	770	Conservative 0	Mismatches 79	Indels 9
				Gaps 1
QY	1	TCAAAATCGAAGAGCTTTAGACTTGTCTTAACCGCCAAAGAGAGGGGGACCTGTTTATTTT	60	
DB	961	TCAAAATCGAAGAGCTTTAGACTTGTCTTAACCGCGAAAGCGGGGAACTCTTTATTTT	1020	
QY	61	AGGGGAAGAAATGCTGTTTAGTATGTTAAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA	120	
DB	1021	AGAGGAAAAATGCTGTTTAGTATGTTAAATCAATCCGGAATCATCACCGAGAAAGTTAAAGA	1080	
QY	121	AATTCGAGATCGAATATATAATGTAGAGAGAGGACCTTCAAAACACTGCACCCCTGGGGCCT	180	
DB	1081	AATTCGAAGTCGAATATAAAGTATAGACAAAGGAGCTGCAAAACACTGGACCCCTGGGGCCT	1140	
QY	181	CCTCAGCCAAATGATGCCCTGGACTCCGCCCTTTAGGACCTCTAGCAGCTATAATATT	240	
DB	1141	CCTCAGCCAAATGATGCCCTGGAAATCTGCCCTTCTTAGGACCTCTAGCAGCTATAATATT	1200	
QY	241	TTTACTCCCTCTTTGGGACCCCTGTATCTTCAACTTCTTGTAGTTTGTCTTCCAGAAT	300	
DB	1201	GTTACTCCCTTTGGACCCCTGTATCTTAACCTCTTGTAGTTTGTCTTCCAGAAT	1260	
QY	301	TGAAGCTGTAAAGCTACAAATAGTCTTCAAAATGGAACCCCGAGTCAGTCCATGACTAA	360	
DB	1261	CGAAGCAGTAAAACTACAAATCGTTTCAAAATGGAGCCCGAGATGCAGTCCATGAGTAA	1320	
QY	361	AATCTACCTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAAATGACATTTGA	420	
DB	1321	AATCTACCAAGGACCCCTGGACCGGCTGCTAGCCATGCTCTGATGTTAAATGACATCAA	1380	
QY	421	AGTCACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG	480	
DB	1381	AGGCAACCCCTCCCGAGGAAATCTCAACTGCACAAACCTCTACTACGCCCAATTCAGCAGG	1440	
QY	481	AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA	540	
DB	1441	AAGCAGTTAGAGTGGTTGTTGGCCAACTCCCAACAGCAGTTCGGTTTTCTGTTGAGA	1500	
QY	541	GGGTGACCTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAGCCT	600	
DB	1501	GGGGGACCTGAGAGACAGGAATAACTAGATTTCTTAGCACTAAGAAATCCCTAGACT	1560	
QY	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCCG	660	
DB	1561	AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCACCTTAGCTCACGCCCA	1620	
QY	661	ACCAATC-----AGAGAGCTCAATAAATGCTTAATCAGGCAAAAACAGAGAGTAAA	711	
DB	1621	ACCAATCAGATCAATAAGAGAGCTCAATAAATGCTTAATTAGGCAAAAACAGAGAGTAAA	1680	
QY	712	GCATAGCCAATCATCTATTGCTCTGAGACACAGCGGAGGACAGGATTTGGGATATAA	771	
DB	1681	GAAATAGCCAATCATCTGTTGCTCTGACACACAGCAGGAGGACAAATGATCGGATATAA	1740	
QY	772	ACTCAGGCATTTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAG	831	
DB	1741	ACCCAGGCATTCGAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGATGGGAG	1800	
QY	832	CTCTGTTTTCACTCTATT 849		
DB	1801	CTCTGCTTCACTCTATT 1818		

DE	Human bone marrow expressed single exon probe SEQ ID NO: 4575.
XX	
XX	Human; bone marrow expressed exon; gene expression analysis; probe;
KW	microarray; cancer; leukaemia; lymphoma; myeloma; ss.
KW	
XX	
OS	Homo sapiens.
XX	
FN	WO200157276-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US000668.
XX	
XX	04-FEB-2000; 2000US-0180312P.
PR	26-MAY-2000; 2000US-0207456P.
PR	30-JUN-2000; 2000US-00608408.
PR	03-AUG-2000; 2000US-00632366.
PR	21-SEP-2000; 2000US-0234687P.
PR	27-SEP-2000; 2000US-0236359P.
PR	04-OCT-2000; 2000GB-00024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
XX	WPI; 2001-488900/53.
DR	
XX	
PT	Human genome-derived single exon nucleic acid probes useful for analyzing
PT	gene expression in human bone marrow.
XX	
PS	Example 4; SEQ ID NO 4575; 658pp + Sequence Listing; English.
XX	
CC	The present invention provides a number of single exon nucleic acid
CC	probes which are derived from genomic sequences expressed in the human
CC	bone marrow. They can be used to measure gene expression in bone marrow
CC	samples, which may enable the improved diagnosis and treatment of cancers
CC	such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC	the probes of the invention
XX	
SO	Sequence 1894 BP: 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

QY		301	TGAAGCTGTTAAAGGTACAAAATAGTTCCTCAAAATGGAAACCCCGACATGCAGTCCTCAACTGACTAA	360
Db		1261	CGAAGCAGTAGTAATACTACAAAATCGTTCTTCAAATGGAGCCCCAGATGCAGTCCAATGAGTAA	1320
QY		361	AATCTACCGTGGACCCTCGACCGCGCTGCTAGACTATGCTCTGTATGTTAATGACATTGA	420
Db		1321	AATCTACACGGAGCCCCTGGACCGGCCCTGCTAGCCCATGCTCTGTATGTTAATGACATCAA	1380

```
QY 421 AGTACCCCTCCCGAGGAAATCTCACTGACACACCCCTACTACACTCCAATTCAGTAGG 480
DB 1381 AGGCACCCCTCCCGAGGAAATCTCACTGACACACCCCTACTACAGCCCAATTCAGCAGG 1440
QY 481 AAGCAGTTAGAGCAGTTCTGACGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
DB 1441 AAGCAGTTAGAGTGGTTGTTGGCCAACTCCCAACAGCAGTTGGGTTTCTGTTGAGA 1500
QY 541 GGGTGGACTGAGAGACAGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCNAAAGCCT 600
DB 1501 GGGGGGACTGAGAGACAGGAATACTAGATTTCCTAGACCAACTAAGAAATCCCTAAGACT 1560
QY 601 ANCTGGGAAGTGACCGCATCATCTTTAAACATGGGCTTGCACACTTAGCTCACACCG 660
DB 1561 AGCTGGGAAGGTGACCGCTTCACCTTTAAACACCGGGCTTGCACACTTAGCTCACGCCCA 1620
QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAA 711
DB 1621 ACCAATCAGACTAAGAGAGCTCACTAATATGCTAATAGGCAAAACAGGAGATAA 1680
QY 712 GCAATAGCCAATCATCTATTGCTGAGAGACAGCGGGAAGACAAAGATTGGGATATAA 771
DB 1681 GAAATAGCCAATCATCTGTTGCTGACAGCACAGCAGGAGGACAAATGATCGGGATATAA 1740
QY 772 ACTCAGGCATTCAGCCGACAAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAG 831
DB 1741 ACCCAGGCATTCGAGCCAGCTACAGTACCCTCTTTGGGTCCCTCCCTTTGATGGGAG 1800
QY 832 CTCTGTTTCACTCTATT 849
DB 1801 CTCTGTTTCACTCTATT 1818
RESULT 34
ID AAK04516
AC AAK04516;
XX
XX 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe SEQ ID NO: 4507.
XX Human; brain expressed exon; gene expression analysis; probe: microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX Homo sapiens.
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000667.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
XX Example 4; SEQ ID NO 4507; 650pp + Sequence Listing; English.
```

```
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
Query Match 53.0%; Score 704.8; DB 4; Length 1894;
Beat Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCAACGGCCAAAGAGGGGAACTGTTTATTTT 60
DB 961 TCAAAATCGAAGAGCTTTAGACTTGTCAACGGCCAAAGAGGGGAACTGTTTATTTT 1020
QY 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
DB 1021 AGAGGAAATGCTGTTGTTATGTTAATCAATCCGGAATCATCACCAGAAAGTTAAAGA 1080
QY 121 AATTGAGATCGAATATAATGTAGAGCAGAGGACCTTCAAAAACACTGACCCCTGGGCT 180
DB 1081 AATTCAAGGTCGAATATAACGTAGAGCAAGAGGAGCTGCAAAAACACTGACCCCTGGGCT 1140
QY 181 CCTCAGCAATGATGCTGCTGGACTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT 240
DB 1141 CCTCAGCAATGATGCTGCTGGACTCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT 1200
QY 241 TTTACTCTCTTTGGAGCCCTGTATCTTTCAACTCTTCTTAAAGTTTCTCTCCAGAA 300
DB 1201 GTTACTCTCTTTGGAGCCCTGTATCTTTAACTCTCTTAAAGTTTCTCTTCCAGAA 1260
QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCGTCCATGACTAA 360
DB 1261 CGAAGCAGTAAGAACTACAAATCGTTCTTCAAAATGGAGCCAGATGCGTCCATGAGTAA 1320
QY 361 AATCTAGCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 1321 AATCTAGCGGACCCCTGGACCGGCTGCTAGCCGATGCTCTGATGTTAATGACATCA 1380
QY 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCAACACCCCTACTACACTCCAATTCAGTAGG 480
DB 1381 AGGCACCCCTCCCGAGGAAATCTCAACTGCAACACCCCTACTACGCCCAATTCAGCAGG 1440
QY 481 AAGCAGTTAGAGCAGTTGTCAGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
DB 1441 AAGCAGTTAGAGTGGTTGTTGGCCAACTCCCAACAGCAGTTGGGTTTCTGTTGAGA 1500
QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCNAAAGCCT 600
DB 1501 GGGGGGACTGAGAGACAGGAATACTAGATTTCCTAGACCAACTAAGAAATCCCTAAGACT 1560
QY 601 ANCTGGGAAGTGACCGCATCCATCTTTAAACATGGGCTTGCACACTTAGCTCACACCG 660
DB 1561 AGCTGGGAAGGTGACCGCTTCACCTTTAAACACCGGGCTTGCACACTTAGCTCACGCCCA 1620
QY 661 ACCAATC-----AGAGAGCTCACTAATGCTAATCAGGCAAAACAGGAGGTAAA 711
DB 1621 ACCAATCAGACTAAGAGAGCTCACTAATATGCTAATAGGCAAAACAGGAGATAA 1680
QY 712 GCAATAGCCAATCATCTATTGCTGAGAGCACAGCGGGAAGGACAAGGATTGGGATATAA 771
DB 1681 GAAATAGCCAATCATCTGTTGCTGACAGCACAGCAGGAGGACAAATGATCGGGATATAA 1740
QY 772 ACTCAGGCATTCAGCCGACAAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAG 831
DB 1741 ACCCAGGCATTCGAGCCAGCTACAGTACCCTCTTTGGGTCCCTCCCTTTGATGGGAG 1800
QY 832 CTCTGTTTCACTCTATT 849
DB 1801 CTCTGTTTCACTCTATT 1818
```


XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX Claim 25; SEQ ID NO 4413; 322pp; English.
XX The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridizes at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;
SQ
Query Match 53.0%; Score 704.8; DB 5; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAATCGCCAAAGAGGGGAACCTGTTATTTTT 60
DB 961 TCAAAATCGAAGAGCTTTAGACTTCTAATCGCCGAAAGCGGGGAACCTTTTATTTTT 1020
QY 61 AGGGAAGAGTCTGTAGTATGTTAAATCAATCTCGAATCATCTACTGAGAAGTTAAAGA 120
DB 1021 AGAGGAAATATCTGTTATGTTATTAATCAATCCGAATCATCCGAGAAGTTAAAGA 1080
QY 121 AATTTGAGATCGAATATAATATAGACGAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
DB 1081 AATTCAGGTCGAATATAACGTFAGACAAAGGAGCTGCAAAACACTGGACCTGGGGCT 1140
QY 181 CCTCAGCAGATGGATGCCCTGAGACTCTGCCCTTCTTAGACCTCTAGCAGCTATAATTT 240
DB 1141 CCTCAGCAGATGGATGCCCTGAGACTCTGCCCTTCTTAGACCTCTAGCAGCTATAATTT 1200
QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTAGTTTGTCTCTTCCAGAT 300
DB 1201 GTTACTCTCTTTGGACCTGATCTTAACTCTCTTGTAGTTTGTCTTTCAGAT 1260
QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTTCAATGGAACCCAGATGCGATCCATCACTAA 360
DB 1261 CGAAGCAGTAAACTACAAATCGTTCTTCAATGAGGCCAGATGCGATCCATGAGTAA 1320
QY 361 AATCTACCTGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 1321 AATCTACCGACCCCTGGACCGGCTCTAGCCATGCTCTGATGTTAATGACATCAA 1380
QY 421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAATTCAGTAG 480
DB 1381 AGGCACCCCTCCCGAGGAATCTCAACTGCACAAACCTCTACTACGCCCAATTCAGCAGG 1440
QY 481 AAGCAGTTAGACAGTTGTGACGCCAACTTCCCAACAGTACTTGGGTTTCTGTTGAGA 540

DB 1441 AAGCAGTTAGAGTGGTTGTTGGCAACCTCCCAACAGCAGTGGGTTTCTCTGTTGAGA 1500
QY 541 GGGTGGACTGAGAGACGAGGACTAGCTGGATTTCTTAGCTGACTAAGAATCCCAAGCCT 600
DB 1501 GGGGGGACTGAGAGACGAGGAAATACTAGATTTCTTAGCCCAACTAAGAATCCCTTAAGACT 1560
QY 601 ANCTGGGAAGGTGACCGCATCTTTAAACATGGGGCTTGGCAACTTAGCTCACACCCG 660
DB 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTCAGGCCA 1620
QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAA 711
DB 1621 ACCAATCAGATACATAAGAGAGCTCACTAAATGCTAATTAGGCAAAAACAGGAGATAA 1680
QY 712 GCATAGCCCAATCATCTATTGCTGAGACGACGCGGGAAGGACAAGGATGGATATAA 771
DB 1681 GAAATAGCCCAATCATCTGTTGCTGACAGCAGCAGGAGGAGCAATGATCGGATATAA 1740
QY 772 ACTCAGGCATTTCAAGCCAGCAACAGCAACCCCTTTTGGGTCCCTCCCATTTGTATGGAG 831
DB 1741 ACCGAGCATTCGAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGTATGGAG 1800
QY 832 CTCTGTTTTCATCTATT 849
DB 1801 CTCTGTTTTCATCTATT 1818
RESULT 37
ABS04589
ID ABS04589 standard; DNA; 1894 BP.
XX
AC ABS04589;
XX
DT 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe from lung SEQ ID No 4580.
DE
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
XX
PN WO200186003-A2.
XX
XX 15-NOV-2001.
PD
XX
PF 30-JAN-2001; 2001WO-US0000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2002-114183/15.
DR
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
XX Claim 1; SEQ ID NO 4580; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1894 BP; 557 A; 500 C; 361 G; 476 T; 0 U; 0 Other;

Query Match 53.0%; Score 704.8; DB 6; Length 1894;
Best Local Similarity 89.7%; Pred. No. 3.3e-220;
Matches 770; Conservative 0; Mismatches 79; Indels 9; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAGAGGGGAACTGTTATTTT 60
DB 961 TCAAAATCGAAGAGCTTTAGACTTGCTAACCGCGGAAAGCGGGGAACTTTTATTTT 1020

QY 61 AGGGGAAGATGCTGTTAGTATGTTTAACTGAATCTGAATCATCTAGAGAAAGTTAAGA 120
DB 1021 AGAGGAAAAATGCTGTTGTTATGTTTAACTCAATCCGGAATCATCCGGAAGAAAGTTAAGA 1080

QY 121 AATTGAGATCGAATATATGATAGACAGAGGACCTTCAAAACACGACCTGGGCGCT 180
DB 1081 AATTCAAGTTCGAATATATGATAGCAAGAGGAGCTGCAAAAACCTGGACCTTGGGCGCT 1140

QY 181 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 240
DB 1141 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 1200

QY 241 TTTACTGCTCTTTGGACCTGATATCTTCAACTCTCTTTGTTAGTTGCTCTTCCGAAT 300
DB 1201 GTTACTGCTCTTTGGACCTGATATCTTCAACTCTCTTTGTTAGTTGCTCTTCCGAAT 1260

QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
DB 1261 CGAAGCAGTAAACTACAAATCGTTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 1320

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
DB 1321 AATCTACCGGACCCCTGGACCGGCTGCTAGCCCATGCTCTGATGTTAATGACATCAA 1380

QY 421 AGTACCCTCCCGAGGAAATCTCACTGCAACACCCCTACTACTCACTCCAATTCAGTAGG 480
DB 1381 AGGACCCCTCCCGAGGAAATCTCACTGCAACACCCCTACTACTAGCCCAATTCAGCAGG 1440

QY 481 AAGCAGTTAGAGCAGTTCCTGAGCCAACTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
DB 1441 AAGCAGTTAGAGTGGTTCTTGGCCCAACCTCCCAACAGCAGTGGGTTTCTCTGTTGAGA 1500

QY 541 GGGTGGACTGAGACAGAGCTAGCTGATTTCTTAGGCTAGTAAAGAAATCCNAAAGCCT 600
DB 1501 GGGGGGACTGAGACAGAGATACTAGATTCTTAGACCACTAAGAAATCCCTTAGACT 1560

QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTTCACACCCG 660
DB 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCAACTTAGCTTCACGCCA 1620

QY 661 ACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGAGGATAA 711
DB 1621 ACCAATCAGATACTAAAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGAGATAA 1680

QY 712 GCATAGCCCAATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTCGGATATAA 771
DB 1681 GAAATAGCCCAATCATCTGTTGCTGACAGCAGCAGGAGGACAAATGATCGGATATAA 1740

QY 772 ACTCAGGCAATTCAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCAATTTGATGGGAG 831
DB 1741 ACCAGGCAATTCAGCCAGCTACAGTACCTCTTTGGTCCCTCCCTTTGATGGGAG 1800

QY 832 CTCTGTTTCACTCTATT 849
DB 1801 CTCTGTTTCACTCTATT 1818

RESULT 38
AAS77313
ID AAS77313 standard; cDNA; 2942 BP.
XX AAS77313;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #13117.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; es.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG13126.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensic, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 13117; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

1982	Db	TGTCACAGACCCACAGCTGACTTCGATCCCTCTG-----	2015
1006	Qy	CTGATCCAGCGAGGTACCAATTGCCACTCCGATCAGGCTAAAGGCTGCCATTGTTCCT	1065
2016	Db	--GATCCAGGAGGCGCCCAATTGCTGCTCCTGATCGGGCTAAAGGCTTGCCCTGTGTTCCT	2073
1066	Qy	GCATGGCTAAGTG-CTTGGGTTTGTCTTAATAGAACTGAAACACTGGTCACTGGTTCCTAT	1124
2074	Db	GCAGGGCTAAGTGCCCGAGGATTCGTCCTAATCTAGCTAAACACTAGTCACTGGGTTCCAC	2133
1125	Qy	GGTTCTCTTCCATGACCCAGCGGCTTCTAATAGAGCTATAACACTCAACCGCATGCCCCAAG	1184
2134	Db	GATTCTCTTCCATGATGACCCAGAGCTTCTAATAGAGCTATAACACTCACTGTCATGCCCCACA	2193
1185	Qy	ATTTCCAATCTCTTGATCTGTGAGGCCAAGAA-CCCCAGGTCAGAGAAANGTGAGCTTGC	1243
2194	Db	GTTCCTAATCTCTTGGAACTCCGTGAGGCCAGAACCCCGAGTTCACAGNAATGAGGCTTGC	2253
1244	Qy	CACCATTTGGGAAGTGGCCCACTGCCATTTTGGTAGGGCCCAACCACTCTTTGGGAGCT	1303
2254	Db	CACCATCTCGAAGAGGCCCGTCCACCATCTTGGAAAGCGACCTTGCCACCACTCTTGGGAAT	2313
1304	Qy	GTGGAGCAAGGATCCCC	1321
2314	Db	TGGGGCAAGGACCCCC	2331

RESULT 39
AAS65964
ID AAS65964 standard; cDNA; 893 BP.
XX
XX AAS65964;
XX
XX 13-FEB-2002 (first entry)
XX
XX 13-ENC-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #1768.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI; 2001-639362/73.
DR
XX
XX P-PSDB; ABG01777.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 1768; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a

polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS61457-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at www.int/pub/published/pct/sequences

SQ Sequence 893 BP; 223 A; 264 C; 202 G; 204 T; 0 U; 0 Other;
 Query Match 49.0%; Score 651; DB 5; Length 893;
 Best Local Similarity 85.0%; Pred. No. 1.1e-202;
 Matches 757; Conservative 0; Mismatches 117; Indels 17; Gaps 2;

Qy	321	TAGTTCCTTCAAATGGAACCCGAGATGAGTCCATGACTAAAAATCTACCGTGGACCCCTGG	380
Db	10	TAAAACTACAAATGGAGCCCAAGATGCAGTCCAAAGACTAAGATCTACTCCGCAGACCCCTGG	69
Qy	381	ACGGCCCTGCTAGACTATGCTCTGATCTTAATGACATTTGAAGTACACCCCTCCCGAGGAAA	440
Db	70	ACGGCCCTGCTAGGCCACGATCTGATGTTAATGACATCAAAAGGCACCCCTCGTGAGGAAA	129
Qy	441	TCTCAACTGCACAAACCCCTACTACACATCCAAATTCAGTAGGAGCAGTTAGAGCAGTTGTC	500
Db	130	TCTCAGCTGCACAACCTCTACTAGCCCCAAATTCAGCAGGAAGCAGTTAGAGCGTCTCT	189
Qy	501	AGCCAACTCCCCAAACAGTACTCTGGGTTTTCTCTTGAGAGGGTGGACTCAGAGACAGGA	560
Db	190	GGCCAACTCCCCAAACAGCATTAAGTTTTCTCTGTTGAGATGGGGACTCAGAGACAGGA	249
Qy	561	CTAGCTGGATTTCTTAGCTGACTTAAGAATCCNAAAGCCTANCTGGGAAAGGTGACCGCAT	620
Db	250	CTAGCTGGATTTCTTAGGCCCACTAAGATCCCTAAGCCTAGCTAGCTGGGAAAGGTAACCAT	309
Qy	621	CCATCTTTTAAACATGGGGCTTGCAACTTAGTGTCAACCCGACCAATC-----AGAG	671
Db	310	CCACCTTTTAAACACGGTGCTTGCAACTCAGCTCAACCCGACCCCATCAGGTAAAGAAAG	369
Qy	672	AGCTCACTTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCCAAATCATCTATT	731
Db	370	AGCCCGCTTAAATGCTAATTAGGCAAAAACAGGAGGTAAGAAATAGCCAAATCATCTATT	429
Qy	732	GCCTGAGAGCACAGCGGGAAGGACAAGGATGGGATATAAACTCAGGCAATTCAGAGCCAGC	791
Db	430	GCCTGAGAGCACAGCGGAGGAGGACATGATCAGATATTAACCCAGGCAATTCGAGCCGGC	489
Qy	792	AACAGCAACCCCTTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTTCACTCATTTTC	851
Db	490	AACGACTACCTCTTTTGGGTCCCTCCCTTTGATGGAGCTCTGTTTTCACTCTATTAA	549
Qy	852	ACTCTATTAAATCATGCAACTGCACTCTTCTGGTCCGHTTTTTTATGCTCAGAGCTGAG	911
Db	550	ATCTTGCAACTGCA-----CACTCTCTGGTCTGTGTTGTTAAGGCTCTGAGCTGAG	601
Qy	912	CTTTTGTGCGCATCCACACTGCTGTTTGGCCACCGTCACAGACCCCGCTGCTGACTTCCA	971
Db	602	CTTTTGGCTCACCGTCCACACTGCTGTTTGGCCACCGTTGCAGACCCCATCGCTGACTTCCA	661
Qy	972	TCCCTTTGGATCCAGCAGAGTCCACTGTGCTCTGATCCAGCGAGGTACCCATTGGCA	1031
Db	662	CCCTTCCAGATATGCGAGGTGTCGCTGCGGTTTCTGATATAGCAAGGCGCCATTGGCG	721
Qy	1032	CTCCGATCAGGCTAAAGGCTTGCCATTGTTTCTTGCAATGGCTAAGTGCCTGGGTGTGTC	1091
Db	722	CTCCCAATCGGGCTAAGGCTCGCCATGTTTCTTGCAATGGCTAAGTGCCTGGGTGTGTC	781
Qy	1092	TAATAGAACTGAACACTGGTCACTGGGTTCATGGTCTCTTTCATGACCCACCGGCTTCT	1151

Db	782	TAATCCAGCTGAACACTAGTTGCTGGGTTCACGGTTCCTTCCATGACCCACGGCTTCT	841
Qy	1152	AATAGAGCTATAAACAACCTCACCGCATGGCCCAAGATTCCATTCTTGGTATC	1202
Db	842	AATAGAGCTATAAACAACCTCACCGCATGGCCCAAGTTCCATTCTTGGAAATC	892
RESULT 40			
ABN97978			
ID	ABN97978	standard; DNA; 46340 BP.	
XX	AC	ABN97978;	
XX	DT	01-AUG-2002 (first entry)	
XX	DE	Human retroviral sequence HI3.	
XX	KW	Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;	
XX	KW	multiple sclerosis; ds.	
XX	OS	Human endogenous retrovirus.	
XX	PN	WO9967395-A1.	
XX	PD	29-DEC-1999.	
XX	PF	23-JUN-1999; 99WO-FR001513.	
XX	PR	23-JUN-1998; 98FR-00007920.	
XX	PA	(INRM) INSERM INST NAT SANTE & RECH MEDICALE.	
XX	PI	Alliel PM, Perin J, Rieger F;	
XX	DR	WPI; 2000-160587/14.	
XX	PT	New nucleic acid sequences of human endogenous retrovirus, HERV-7q, used	
XX	PT	for diagnosis, treatment and prevention of autoimmune and neurological	
XX	PS	Claim 15; Page 186-199; 225pp; French.	
XX	CC	The present invention relates to new nucleic acid sequences of human	
XX	CC	endogenous retrovirus, HERV-7q, which is located on chromosome 7q.	
XX	CC	Regulatory elements associated with HERV-7q may alter expression of other	
XX	CC	genes (even remote genes) on the same chromosome, inducing immunological	
XX	CC	and/or neurological changes (which may be pathological or protective/	
XX	CC	curative). HERV-7q peptides can be used to improve efficiency of the	
XX	CC	immune response, e.g. in immunotherapy. HERV-7q peptides and their coding	
XX	CC	sequences can be used in immunogenic or vaccinating compositions, for	
XX	CC	protection against autoimmune diseases, particularly multiple sclerosis.	
XX	CC	The peptides may also be used (by sequence comparison) to detect/identify	
XX	CC	endogenous retroviruses that are abnormally expressed in cancer,	
XX	CC	neuropathologies or other autoimmune diseases. The present sequence was	
XX	CC	used to illustrate the invention	
XX	SQ	Sequence 46340 BP; 16104 A; 8738 C; 8434 G; 13064 T; 0 U; 0 Other;	
	Query Match	48.4%; Score 643.4; DB 3; Length 46340;	
	Best Local Similarity	85.4%; Pred. No. 3.6e-199;	
	Matches 781; Conservative	0; Mismatches 88; Indels 46; Gaps 4;	
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGCTAACCCGCAAAAGAGGGGAACCTGTTTATTTT 60	
Db	32993	TCAAAATCGAAGAGCTTTAGACTTGCTAACCTCGAAAGAGGGGAAGCTTTTATTTT 33052	
Qy	61	AGGGGAAGAAATGCTGTAG-----TATGTTAATCA 90	
Db	33053	AGGGGAAGAAATGTTGTTTATGTTATTTTAGCGGAAGAAATGTTGTTTATGTTAATCA 33112	
Qy	91	ATCTGGAATCATTTACTGAGAAAGCTTAAAGAAATTTGAGATCGCAATAATAATGAGAGAGA 150	

Fri Feb 25 16:26:28 2005

us-09-319-156b-12.rng

Db 33113 ATCCTGAATTCTCACAGAAAAGTTGAAGAAATTCGAGATTGATATACAACGTAGAACAGA 33172
Qy 151 GGAACCTTC-AAAACACTGCACCCCTGGGGCCCTCCTCAGCCAATGGATGCCCTGGACTCTCC 209
Db 33173 GGAGCTTCAAAAACACAGACCCCTGGGGCCCTCCTCAGCCAATGGATGCCCTGGATCTCC 33232
Qy 210 CCTTCTTAGGACCTCTAGCAGCTATAATATTTTTACTCCTCTTTTGGACCCCTGTATCTTCA 269
Db 33233 CCTTCTTAGGATCTTAGCAGCTCTAATATTTGATACTCCTCTTTGGACCCCTGTATCTTTA 33292
Qy 270 ACTTCTTTGTTAAGTTTGTCTCTTCCAGAAATTGAAGCTGTAAAGCTACAAATAGTTCTTTC 329
Db 33293 ACCTCCTTTGTTAAGTTTGTCTCTTCCAGAAATCAAAAGTTGTAAAGCTACAAATCGTTCTTC 33352
Qy 330 AAATGGRAACCCAGATGCAGTCCATGACTAAAAATCTACCGTGGACCCCTGGACCCGGCTG 389
Db 33353 AAATGGAAACCCAGATGAAGTCCATGACTAAGATCTACCGTGGACCCCTGGACCCGGCTA 33412
Qy 390 CTAGACTATGCTCTGAATGTTAATGACATTGAAGTCAACCCCTCCCGAGGAAATCTCAACTG 449
Db 33413 CTAGCCCATGCTCCAAATTTGTAATGATATCGAAACGACCCCTCCCGAGGAAATCTCAACTG 33472
Qy 450 CACAACCCCTACTACACTCCAAATTCAGTAGGAGCAGTTAGAGCAGTTGTCTCAGCCAACT 509
Db 33473 CACAACCCCTACTATGCCCCAAATCCCGAGGAAGCAGTTAGACTGCTCTCAGCCAACT 33532
Qy 510 CCCCAACAGTACTTGGGTTTCTGTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGA 569
Db 33533 CCCCAACAGCACTTGGGTTTCTGTTGAGTGGGGGACTGAGAGACAGGATTAGCTGA 33592
Qy 570 TTTCCTAGGCTGACTAAGAAATCCCNAACTTCCGAAAGTGGGAAGTGACCGCATCCATCTTTA 629
Db 33593 TTTCCTAGGCCGACTAAGAAATCCCAAGCCTAGCTGGGAAGTGACCCACATCCACCTTTA 33652
Qy 630 AACATGGGGCTTGGCAACTTAGCTCACACCCGACCAATC-----AGAGAGCTCACTA 680
Db 33653 AACATGGGGCTTGGCAACTTAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCTCACTA 33712
Qy 681 AAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCCAATCATCTATTGCTGAGAG 740
Db 33713 AAATGCTAATAGACAAAACAGGAGGTAAGCAATAGCCAATCATCTATCGCTGAGAG 33772
Qy 741 CACAGCGGAGGACAAAGGATTTGGGATATAAACTCAGGCATTTCAAGCCAGCAACAGCAAC 800
Db 33773 CACAGCGGAGGACAAATGATCGGGATATAAAACCCAGGCATTTCAAGCCGCAACGCTAC 33832
Qy 801 CCCCTTTGGGTCCCTCCCTTGTATGGGAGCTCTGTTTCACTCTATTCTCTATTATA 860
Db 33833 CTTCCTTTGGGTCCCTCCCTTGTATGGGAGCTCTCTCT-----GTCTTCACTCTATTATA 33886
Qy 861 AATCATGCAACTGCA 875
Db 33887 AATATTGCAACTGCA 33901

Search completed: February 21, 2005, 07:46:02
Job time : 696.525 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:54:34 : Search time 234.166 Seconds
(without alignments)
9286.612 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatcgagagcttag.....gcaagatccccagtaaca 1329

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	1326	99.8	1329	4	US-08-979-847B-108
2	1035.2	77.9	13537	4	Sequence 17382, A
3	731.8	55.1	2946	3	Sequence 3, Appli
4	729.2	54.9	2763	4	Sequence 5640, Ap
C 5	728.4	54.8	168394	4	US-09-949-016-13002
C 6	654.8	49.3	77772	4	US-09-949-016-17417
C 7	654.8	49.3	77797	4	US-09-949-016-12249
C 8	643.4	48.4	145320	4	US-09-949-016-15858
C 9	637.8	48.0	780	4	Sequence 385, App
C 10	634.6	47.8	245286	4	US-09-949-016-15497
C 11	582	43.8	219964	4	US-09-949-016-15086
C 12	565.6	42.6	84571	4	US-09-949-016-17420
C 13	562	42.3	194790	4	US-09-949-016-15393
C 14	545	41.0	635	4	US-08-979-847B-102
C 15	500.2	37.6	140234	4	US-09-949-016-17002
C 16	478.6	36.0	116552	4	Sequence 13413, A
C 17	468.6	35.3	131631	4	US-09-949-016-11757
C 18	467.2	35.2	89584	4	US-09-949-016-17068
C 19	446	33.6	154023	4	US-09-949-016-17057
C 20	445.2	33.5	251672	4	US-09-949-016-17296
C 21	445.2	33.5	251682	4	US-09-949-016-11973
C 22	440.4	33.1	8523	4	US-09-949-016-17002
C 23	435.8	32.8	99580	4	US-09-949-016-17411
C 24	434.8	32.7	601	4	US-09-949-016-202370
C 25	427.8	32.2	601	4	US-09-949-016-44299
C 26	411.4	31.0	57507	4	US-09-949-016-15019
C 27	399.8	30.1	39686	4	US-09-949-016-13633

C 28	399.8	30.1	49487	4	US-09-949-016-11770
C 29	387.4	29.1	601	4	US-09-949-016-44300
C 30	371	27.9	601	4	US-09-949-016-195793
C 31	369.6	27.8	601	4	US-09-949-016-44298
C 32	345.4	26.0	601	4	US-09-949-016-146564
C 33	344.6	25.9	443	4	US-09-513-999C-21394
C 34	343.6	25.9	149971	4	US-09-949-016-13590
C 35	338.8	25.5	601	4	US-09-949-016-133739
C 36	336	25.3	601	4	US-09-949-016-120277
C 37	324.6	24.4	601	4	US-09-949-016-202257
C 38	317.8	23.9	601	4	US-09-949-016-146565
C 39	317.6	23.9	601	4	US-09-949-016-44297
C 40	316.4	23.8	601	4	US-09-949-016-44296
C 41	311.6	23.4	601	4	US-09-949-016-195792
C 42	305.4	23.0	601	4	US-09-949-016-183390
C 43	302	22.7	1481	4	US-08-979-847B-105
C 44	273.6	20.6	601	4	US-09-949-016-195794
C 45	268.4	20.2	601	4	US-09-949-016-44310

ALIGNMENTS

RESULT 1

US-08-979-847B-108
Sequence 108, Application US/08979847B
Patent No. 6582703

GENERAL INFORMATION:

APPLICANT: PERRON, HERVE
BESEME, FREDERIC
BEDIN, FREDERIC
PARAMHOS-BACALA, GLAUCIA
KOMURIAN-PRADEL, FLORENCE
JOLIVET-REYNAUD, COLETTE
MANDRAND, BERNARD
GARSON, JEREMY
TUKER, PHILIP

TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
THERAPEUTIC PURPOSES

NUMBER OF SEQUENCES: 210

CORRESPONDENCE ADDRESS:

ADDRESSEE: OLIFF & BERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
ZIP: 22320

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/979,847B

FILING DATE: 26-No. 6582703-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 108:

SEQUENCE CHARACTERISTICS:

LENGTH: 1329 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 108:

US-08-979-847B-108

Query Match									
Best Local Similarity 99.8%; Score 1326; DB 4; Length 1329;									
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTT	60						
Db	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTT	60						
Qy	61	AGGGGAAGAAATGCTGTAGTATGTAAATCAATCTGGAATCAATTAAGTCTGAGAAAGTTAAAGA	120						
Db	61	AGGGGAAGAAATGCTGTAGTATGTAAATCAATCTGGAATCAATTAAGTCTGAGAAAGTTAAAGA	120						
Qy	121	AATTGAGATCGAATATAATATAGAGCAGAGAGACTTCAAAACACCTGCAACCTCGGGCCT	180						
Db	121	AATTGAGATCGAATATAATATAGAGCAGAGAGACTTCAAAACACCTGCAACCTCGGGCCT	180						
Qy	181	CCTCAGCCAAATGGATGCCCTGACCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT	240						
Db	181	CCTCAGCCAAATGGATGCCCTGACCTCCCTCTTCTAGGACCTCTAGCAGCTATAATTT	240						
Qy	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTTGTCTCTCCAGAAAT	300						
Db	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTTGTCTCTCCAGAAAT	300						
Qy	301	TGAAGCTGAAGCTACAAATAGTTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA	360						
Db	301	TGAAGCTGAAGCTACAAATAGTTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA	360						
Qy	361	AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA	420						
Db	361	AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA	420						
Qy	421	AGTCAACCCCTCCGAGGAAATCTCAACTGSCAACCCCTACTACACTCCAATCAGTAGG	480						
Db	421	AGTCAACCCCTCCGAGGAAATCTCAACTGSCAACCCCTACTACACTCCAATCAGTAGG	480						
Qy	481	AAGCAGTTAGAGCAGTTGTGACCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA	540						
Db	481	AAGCAGTTAGAGCAGTTGTGACCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA	540						
Qy	541	GGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT	600						
Db	541	GGTGGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT	600						
Qy	601	ANCTGGGAAGTGAACCGCATCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	660						
Db	601	ANCTGGGAAGTGAACCGCATCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG	660						
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720						
Db	661	ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAGCAATAGCC	720						
Qy	721	AATCATCTATTCCCTCAGAGCACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780						
Db	721	AATCATCTATTCCCTCAGAGCACAGCGGGAAGGACAAGGATTTGGGATATAAATCAGGCA	780						
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	840						
Db	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTT	840						
Qy	841	CACTCTATTCTACTATTAAATCATGCAACTGCACTCTCTGTCGGTCCGTTTTTATGG	900						
Db	841	CACTCTATTCTACTATTAAATCATGCAACTGCACTCTCTGTCGGTCCGTTTTTATGG	900						
Qy	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCT	960						
Db	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCGCT	960						
Qy	961	GCTGACTTCCATCCCTTTGGATCCAGAGAGTGTCCATGTGCTCTGATCCAGGAGGT	1020						
Db	961	GCTGACTTCCATCCCTTTGGATCCAGAGAGTGTCCATGTGCTCTGATCCAGGAGGT	1020						

Qy	1021	ACCAATGGCACTCCGATCAGGCTAAAGGCTTGCCATTTGTTCTGTCATGGCTAAGTGCC	1080						
Db	1021	ACCAATGGCACTCCGATCAGGCTAAAGGCTTGCCATTTGTTCTGTCATGGCTAAGTGCC	1080						
Qy	1081	TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTTCCATGAC	1140						
Db	1081	TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTTCCATGAC	1140						
Qy	1141	CCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGTA	1200						
Db	1141	CCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCCTTGTA	1200						
Qy	1201	TCGTGAGGCCAAGAACCCAGGTTCAGAGAAAGTGAGGCTTGCCACCATTTGGGAAGTGG	1260						
Db	1201	TCGTGAGGCCAAGAACCCAGGTTCAGAGAAAGTGAGGCTTGCCACCATTTGGGAAGTGG	1260						
Qy	1261	CCCACTGCCCCATTTTGGTAGCGGCCACCACTCTTTGGGAGCTGTGGAGCAAGGATCCC	1320						
Db	1261	CCCACTGCCCCATTTTGGTAGCGGCCACCACTCTTTGGGAGCTGTGGAGCAAGGATCCC	1320						
Qy	1321	CCAGTAACA 1329							
Db	1321	CCAGTAACA 1329							

RESULT 2

US-09-949-016-17382
; Sequence 17382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/0949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17382
; LENGTH: 13537
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17382

Query Match 77.9%; Score 1035.2; DB 4; Length 13537;									
Best Local Similarity 90.1%; Pred. No. 0;									
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;									
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGGGGGAACCTGTTTATTTT	60						
Db	10675	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTT	10734						
Qy	61	AGGGGAAGAAATGCTGTTAGTATGTTAATCAATCTGGAATCAATTAAGTGAAGTTAAAGA	120						
Db	10735	AGGGGAAGAAATGCTGTTATTTATGTTAATCAATCTGGAATCTGTAAGAAAGTTAAAGA	10794						
Qy	121	AATTTGAGATCGAATAATAATGTAGAGCAGAGGACCTTCAAAACACTGCACCTCGGGCCT	180						
Db	10795	AATTCGAGATCGAATAACAACGTAGAGCAGAGGAGCTTCGAAACACTGGACCTCGGGCCT	10854						
Qy	181	CCTCAGCCAAATGGATGCGCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT	240						
Db	10855	CCTCAGCCAAATGGATGCGCTGGATTTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT	10914						
Qy	241	TTTACTCTCTTTGAGCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT	300						
Db	10915	GCTACTCTCTTTGAGCCCTGTATCTTAACTCTTGTAACTTTGTCTCTTCCAGAAAT	10974						

QY	301	TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATGCGAGTCCATGACTAA	360	APPLICANT: LaVallie, Edward R.
DB	10975	CGAAGCTGTAAACTA-----CAAAATGGAGCCAAAGATGCGAGTCCAAAGCTAA	11022	APPLICANT: Collins-Racie, Lisa A.
QY	361	AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA	420	APPLICANT: Evans, Cheryl
DB	11023	GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCA	11082	APPLICANT: Merberg, David
QY	421	AGTCAACCCCTCCCGAGGAATCTCAATGTCACAACCCCTTACTACCTCAATTCAGTAGG	480	APPLICANT: Treacy, Maurice
DB	11083	AGGCAACCCCTCTCTAGGAATCTCAGCTGTCACAACCTTACTACGCCCCCAATTCAGCAG	11142	APPLICANT: Mi, Sha
QY	481	AGCAGTTAGACAGTTGTGCGCCACCTCCCAACAGTACTTGGTCTTCTGTTGAGA	540	APPLICANT: Genetics Institute, Inc.
DB	11143	AAGCAGTTAGACGGTGTGCGCCCAACCTCCCAACAGCAGTACTTGGTCTTCTGTTGAGA	11202	TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
QY	541	GGTGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCAAGCCT	600	FILE REFERENCE: 6006B AJ172A
DB	11203	TGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCTAAGCCT	11262	CURRENT APPLICATION NUMBER: US/09/175,928A
QY	601	ANCTGGGAAGTGACCGCATCCATCTTTAAACAATGGGGCTTGCAACTTAGCTCACACCG	660	CURRENT FILING DATE: 1998-10-20
DB	11263	AGCTGGGAAGTGACCACTCACTTTAAACAACGGGGCTTGCAACTTAGCTCACACCTG	11322	NUMBER OF SEQ ID NOS: 62
QY	661	ACCAATCAGAGAGCTCACTAAATGCTANTCAGGCAAAACAGGAGGTAAAGCAATAGCC	720	SOFTWARE: PatentIn Ver. 2.0
DB	11323	ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGCAATAGCC	11382	SEQ ID NO 3
QY	721	AATCATCTAATTCCTGAGAGCAGCGGAAGCAAGGATTTGGGATATAAACTCAGGCA	780	LENGTH: 2946
DB	11383	AATCATCTAATTCCTGAGAGCAGCGGAAGCAAGGATTTGGGATATAAACTCAGGCA	11442	TYPE: DNA
QY	781	TTCAAGCCAGCAACCAACCCCTTTGGGTCCCTTCCATTTGTATGGAGCTCTGTTTT	840	ORGANISM: Homo sapiens
DB	11443	TTCGAGCGCAAGCGCAACCCCTTTGGGTCCCTTCCATTTGTATGGAGCTCTGTTTT	11502	US-09-175-928-3
QY	841	CACTCTAATTTCACTTAAATCATGCAACTGCACTCTTCTGGTCCGTTTATGG	900	Query Match 55.1%; Score 731.8; DB 3; Length 2946;
DB	11503	CATGCTAATTTCACTTAAATCATGCAACTGCACTCTTCTGGTCCGTTTATGG	960	Best Local Similarity 90.7%; Pred. No. 9,1e-240;
QY	901	CTCAAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCCGTCACAGACCCGCT	11622	Matches 794; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
DB	11563	CTTGAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCCGTCACAGACCCGCT	1020	
QY	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCATGCTGCTCTGATCCAGCAGGT	11682	
DB	11623	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCATGCTGCTCTGATCCAGCAGGT	1080	
QY	1021	ACCAATGCACTCCCGATCAGGCTAAGGCTTGCATTTGCTTCGATGGCTTAAGTGCC	11742	
DB	11683	ACCAATGCGGCTCCCAATCGGGCTAAGGCTTGCATTTGCTTCGATGGCTTAAGTGCC	1140	
QY	1081	TGGGTTTGTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTTCCATGAC	11802	
DB	11743	TGGGTTTGTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTTCTCTGAC	1200	
QY	1141	CCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA	11861	
DB	11803	CCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCATTCCTTGGTA	1257	
QY	1201	TCGTGAGGCCAAGAACCCAGCTCAGAGNAGTGGCTTGCACCAATTTGGGAAG	11918	
DB	11862	TCCATAAGGCCAAGAACCCAGCTCAGAGNAGTGGCTTGCACCAATTTGGGAAG		

RESULT 3
US-09-175-928-3
; Sequence 3, Application US/09175928A
; Patent No. 6312921
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.

QY	1023	CCATTGCCACTCCGATCAGGCTAAAGGCTTGCCATTGTTCTCGATGGCTAAGTGCTG	1082
DB	15201	CGGTTGCTCTCCCG---GGCTAGAGACTCGCCATTGTTCTGTCACAGCTAAGTGCTG	15146
QY	1083	GGTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTCTCCATGACCC	1142
DB	15145	GGTTCACTCTTAATTGAGGTGAACACTAGTCGCTGGGTTCCACAGTTCTCTTCCGTGACCC	15086
QY	1143	ACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCTTTGGTATC	1202
DB	15085	ACAGCTTCTTAATAGACTATAACACTCACCGCATGGCCNAGGTTCCATTCTTTGGAATC	15026
QY	1203	TGTGAGGCCAAGAACCCAGGTCAGAGAANTGTAGGCTTTGCCACAATTGG	1253
DB	15025	CATGAGGCTCAGAACCTCCAGGTCAGAGACGAAGACCTTGCTGCCATTGGG	14975

RESULT 7

```

US-09-949-016-12249/c
; RESULT: /
; Sequence 12249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,769
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

```

Query Match 49.3% Score 654.8 DB 4: length 77997:

	Query Match	49.3%	Score 654.8	DB 4	Length 77997	
	Best Local Similarity	83.5%	Pred. No. 2.5e-212			
	Matches 844	Conservative	0	Mismatches 140	Indels 27	Gaps 8
Qy	257	CCCTGATATCTTCAACTTCCCTTGTAAAGTTTGTCCTCTCCAGAAATTCGAAGCTGTAAAGCTA	316			
Db	15976					
Qy	317	CAAAATAGTCTTTCAAATGGAAACCCAGATGCAGTCCATGACTTAAATCTACCGTGGACCC	376			
Db	15916	CAAAAT-GTTCTTCAAATGGAGCCCAAGATGCAGTCCATGACTTAAAGATCTACCGTGGACCC				
Qy	377	CTGGACCGGCTGCTAGACTATGCTCTGATGTTTAAATGACATTTGAATGTCACCCCTCCCGAG	436			
Db	15857	CTGGACTGGCTGCTAGCCCATGCTCCAAATGTTTGATGACATCGAAGGCCCACTCCCGAG	15798			
Qy	437	GAATCTCAATGTGACAACCCCTTACTACCTCCAAATTCAGTAGGAGCAGTTAGAGCAGT	496			
Db	15797	GAATCTCAATGTGATGACCTCTACTACGCCCAATTCAGCAGAGAGCAGTTACAGCGT	15738			
Qy	497	TGTCAGCAACCTCCCAACAGTACTTGGGTTTTTCTGTGTGAGAGGCTGGACTGAGAGAC	556			
Db	15737	CATCGGCCAACCTCCCAACAGCACTTGGGTTTTTCTGTGTGAGAGGGGTACTTAAGAGAC	15678			
Qy	557	AGGACTAGCTGGATTTTCTTAGGCTTGACTAAGAATCCCNAAAGCCTANCT-GGGAAAGTGAC	615			
Db	15677	AGGACTAGCTGGATTTTCTTAGGCTTGACTAAGAATCCCNAAAGCCTANCT-GGGAAAGTGAC	15618			

Qy	616	CGCATCCATCTTTAAACATATGGGGCTTGCAACTTAGCTCACACCCGAGCAATC-----	667
Db	15617	CACACCCACCTTTAAACATATGGGGCTTGTAACCTAGCTCACACCCAGACCAATCAGGTAGTA	15558
Qy	668	-AGAGAGCTCACTAAATGCTAATCAGGC-AAAAACAGGAGGTAAAGCAATAGCCA-ATC	724
Db	15557	AAGAGAGCTCACTAAATACCAATGAGGCTAAANAACAGGTGGTAAGAATAATCAGATC	15498
Qy	725	ATCTATTGCCTGAGAGCACAGCGGGAAGGACAAGGATTTGGGATATAAACTCAGGCATTTCA	784
Db	15497	ATCTGTGCGCTGAGAGCACAGGGGAGGGACAATGATCAGGATATAAAACCCAGGCAATTTCA	15438
Qy	785	AGCCAGCAACAGCAACCCCTTTGGTGCCCTCCCATTTGATGAGAGCTCTGTTTTCACT	844
Db	15437	AGCCGAATGGGGCAACCCCTTTGGGTCCCTCCCGTTGCAATGGAGCTCTGTTTTACG	15378
Qy	845	CTATTTCACCTCTATTAAATCATGCAACTGCACTCTTCTGTGTCCTGTTTTTATGGCTCA	904
Db	15377	ATATTAAATCTTGCAACTGCA-----CACTTTTCTGGTCCATGTTGTTGTCGGTTTG	15326
Qy	905	AGCTGAGCTTTTGTGTGGCCATCCACCACTGCTGTTTTGBCACCGTCCACAGACCCCTGCTG	964
Db	15325	AGCTGAGCTTTTGTGTGGCCATCCACCACTGCTGATGGCTGTGTCGCGAGACGCCCGCTG	15266
Qy	965	ACTTCCA--TCCCTTTGGATCCAGCAGAGTGTCCACTGTCTCTGTATCCACGCGAGGTAC	1022
Db	15265	ACTTCCACCCCTCCAGATCCAGCAGAGGTGTCCACTGCGCTTCTGATCCAGCGAGGCAC	15206
Qy	1023	CCATTGCCACTCCGATCAGGCTAAAGGCTGGCCATTGTTCTGTGATGGCTAAGTGCCCTG	1082
Db	15205	CGGTTGCTGCTCCCG---GGCTAGAGACTCGCCATTGTTCTGTCGACAGCTAAGTGCCCTG	15150
Qy	1083	GGTTGTCTCTAATAGAACTGAACACTGTGTCACTGGTTCCATGGTTCTCTCCATGACCC	1142
Db	15149	GGTTCACTCTAATGAGGTGAACACTAGTGTGCTGGGTTCCACAGTTCTTTCCGTGACCC	15090
Qy	1143	ACGGCTCTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCAATCTTTGGTATC	1202
Db	15089	ACAGCTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGTTCCATCTCTTGGAAATC	15030
Qy	1203	TGTAGGCCAAGAAACCCAGGTCAGAGAANGTGAGGCTTGCCACCAATTTGG	1253
Db	15029	CATGAGGCCAAGAAACCCAGGTCAGAGAGCAAGAGCCTTGTCGCAATTTGGGGTTGG	14979

RESULT 8

```

US-09-949-016-15858/c
; Sequence 15858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15858
; LENGTH: 145320
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15858

Query Match          48.4%;      Score 643.4;  DB 4;      Length 145320;
Best Local Similarity 85.4%;      Pred. No. 3.2e-208;

```

	Matches	781;	Conservative	0;	Mismatches	88;	Indels	46;	Gaps	4;
Qy	1	TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAAGAGGGGAACCTGTTTATTTTT	60							
Db	92888	TCAAAATTGAAAGCTTTAGACTTCTAAACCTCTGAAAGAGGGGAAGCTGTTTATTTTT	92829							
Qy	61	AGGGGAAGAATGCTGTTAG-----TAATGTTAATCA	90							
Db	92828	AGGGGAAGAATGTTGTTATTTATGTTATTTTAGCGGAAGAATGTTGTTATTTATGTTAATCA	92769							
Qy	91	ATCTGGAATCATTTACTTGAGAAAGTTTAAAGAAATTTGAGATCGAATATATATCTAGAGCAGA	150							
Db	92768	ATCCTGAATTTGTCACAGAGAAAGTTGAGAAATTCAGATTTGAATACAACTGAGACAGA	92709							
Qy	151	GGACCTTC-AAAACACTGCACCTCGGGCCTCCTCAGCCAAATGGATCGCCTGGACTCTCC	209							
Db	92708	GGAGCTTCAAAAACACAGACCTCGGGCCTCCTCAGCCAAATGGATCGCCTGGATTTCTCC	92649							
Qy	210	CCCTCTTAGGACCTCTAGCAGCTATAATTTTTTACTCTCTTTTGGACCCCTGATCTTCA	269							
Db	92648	CCCTCTTAGGATCTCTAGCAGCTCAATATTTGATACTCCTCTTTTGGACCCCTGATCTTTA	92589							
Qy	270	ACTTCTCTGTTAAGTTTGCTCTTCCAGAAATTTGAAGCTGTAAGCTCAAAATAGTCTTCTC	329							
Db	92588	ACCTCTCTGTTAAGTTTGCTCTTCCAGAAATCAAAGTTGTAAGCTCAAAATCGTCTTCTC	92529							
Qy	330	AAATGGAACCCAGATGCACTCCATGACTAAAATCTACCGTGGACCCCTGGACCGGCCTG	389							
Db	92528	AAATGGAACCCAGATGAAGTCCATGACTAAGATCTACCGTGGACCCCTGGACCGGCCTA	92469							
Qy	390	CTAGACTATGCTCTGATGTTAATGACTTGAAGTCAACCCCTCCGAGGAAATCTCAACTG	449							
Db	92468	CTAGCCCATGCTCCAATTTGTAATGATATCGAACGCAACCCCTCCGAGGAAATCTCAACTG	92409							
Qy	450	CACAAACCCCTACTACACTCCAATTTAGTAGGAGCAGTTAGAGCAGTTGTCAGCCCAACT	509							
Db	92408	CACAAACCCCTACTATGCCCAATTCGCGAGGAAGCAGTTAGACTGTCGTCAGCCCAACT	92349							
Qy	510	CCCCAACAGTACTTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGA	569							
Db	92348	CCCCAACAGCACTTGGGTTTTCTGTTGAGTGGGGGACTGAGAGACAGGATTTAGCTGGA	92289							
Qy	570	TTTTCTTAGGCTGACTAAGATCCNAGCCTTANTCTGGGAAGGTGACCGCATCATCTTTTA	629							
Db	92288	TTTTCTTAGGCGGACTAAGAAATCCAAAGCCTAGCTGGGAAGGTGACCAATCCACTTTTA	92229							
Qy	630	AACATGGGGCTTGCAACTTAGCTCACACCCGACCAATC-----AGAGAGCTCACTA	680							
Db	92228	AACATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCTCACTA	92169							
Qy	681	AAATGCTAAATCAGGCAAAAACAGAGAGGTAAAGCAATAGCCATCATTTTCCTGAGAG	740							
Db	92168	AAATGCTAAATAGACAAAAACAGAGAGGTAAAAATAGCCATCATCTATCCCTCTGAGAG	92109							
Qy	741	CACAGCGGGAAGGACAGGATTCGGATATAAATCTCAGGCAATTCAGCCAGCAACAGCAAC	800							
Db	92108	CACAGCGGGAAGGACAGGATTCGGATATAAATCCAGGCAATTCAGCCAGCAACAGGCTAC	92049							
Qy	801	CCCCCTTTGGGTCCCTCCCAATGATGGGAGCTCTGTTTTCATCTTATTTCACTCTATTTA	860							
Db	92048	CTTCTTTGGGTCCCTCCCTTTGATGGGAGCTCTCTCT-----GTCCTTCACTCTATTA	91995							
Qy	861	AATCATGCAACTGCA	875							
Db	91994	AATATTGCAACTGCA	91980							

RESULT 9
US-09-573-080A-385
; Sequence 385, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOWL

```

; APPLICANT:  ROGAN, PETER
; TITLE OF INVENTION:  SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE:  30307
; CURRENT APPLICATION NUMBER:  US/09/573,080A
; CURRENT FILING DATE:  2000-05-16
; NUMBER OF SEQ ID NOS:  479
; SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 385
; LENGTH:  780
; TYPE:  DNA
; ORGANISM:  Homo sapiens
; FEATURE:
; NAME/KEY:  repeat region
; LOCATION:  (1)..(780)
; OTHER INFORMATION:  ltr17
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:  Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE:  Prototypic sequences for human repetitive DNA
; JOURNAL:  Journal of Molecular Evolution
; VOLUME:  35
; ISSUE:  4
; PAGES:  286-291
; DATE:  1992-10-
; DATABASE ACCESSION NUMBER:  Database of repetitive elements (repbase)
; DATABASE ENTRY DATE:
; DATABASE ENTRY DATE:  1996-01-26
; US-09-573-080A-385

```

Query Match	48.0%	Score 637.8	DB 4	Length 780
Best Local Similarity	90.0%	Pred. No. 6.8e-208		
Matches 711; Conservative	0	Mismatches 60	Indels 19	Gaps 2
Qy	549	TGAGAGACGAGCTAGCTGGATTTCTCTAGGCTGACTAAGAATCCNAAAGCCTANCTGGGA	608	
Db	1	TGAGAGACAGACTAGCTGGATTTCTCTAGGCGCACTAAGAATCCTTAAGCCTAGCTGGGA	60	
Qy	609	AGGTGACCGCATCCATCTTTTAAACATGGGGCTTGGCACTTTAGCTTCACACCGGACCAATC-	667	
Db	61	AGGTGACCGCTTCACCTTTTAAACACGGGGCTTGGCACTTTAGCTTCACACCGGACCAATCA	120	
Qy	668	-----ACGAGAGCTCACATAAATGCTAATCAGGCAAAAAACGAGAGGTAAAGCAATAGC	719	
Db	121	GGTAGTAAAGAGAGCTCCTATAAATGCTAATTAGGCAAAAAACGAGAGGTAAAGHAATAGC	180	
Qy	720	CAATCATCTATTGCCTGAGAGCACAGCGGAAAGACAAGGATTGGGATATAAACTCAGGC	779	
Db	181	CAATCATCTATTGCCTGAGAGCACAGCGGAGGAGACAATGATCGGATATAAAACCCAGGC	240	
Qy	780	ATTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCACTCTGATGGGAGCTCTGTTT	839	
Db	241	ATTGAGCCCGCAACAGGCTACCTCTTTGGGTCCCTCCCTTTGTATGGGAGCTCTGT--	298	
Qy	840	TCACCTCTATTTCACCTCTATTAAATCATGCACCTGCTCTCTGGTCCGTGTTTTTATG	899	
Db	299	-----TTTCACCTCTATTAAATCTTGAACCTGCACTCTTCTGGTCCGTGTTGTATG	350	
Qy	900	GCTCAAGCTGAGCTTTTGTGCGCATCCACACACTGTGTTTGGCCACCGTCAAGACCCGC	959	
Db	351	GCTCGAGCTGAGCTTTGCTGCTCACCGTCCACCACTGCTGTTTGGCCCGGTGCGAGACCCGC	410	
Qy	960	TGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTGATCCAGCGAGG	1015	
Db	411	CGCTGACTTTCATCCCTCCGGATCCGGCAGGAGTGTCCGCTGTGCTCTGATCCAGCGAGG	470	
Qy	1020	TACCCATTGCCACTCCCGATCAGGCTAAAGGCTTTGCCATTTGTTCTGTGATGGCTAAGTGC	1075	
Db	471	CGCCCAATTGCCGCTCCCGATCGGGCTAAAGGCTTTGCCATTTGTTCTGTGACGCGCTAAGTGC	530	
Qy	1080	CTGGGTTTGCTCTAATAGAACTGAACACTGGTCTACTGGTTCCATGGTTCTCTTCCATGA	1133	
Db	531	CTGGGTTTGCTCTAATCGAGCTGAACACTAGTCACTGGGTTTCCACGGTTTTTTCCCGTGA	590	

Db	471	CGCCCAITGCGCTCCCGATCGGGCTAAAGGCTTGGCAATTGTTCTGTGCACGGCTAAAGTGC	533
Qy	1080	CTGGGTTTGCTTAATAGAACTGAACACTGGTCACTGGTTCCATGGTTCTCTCTCCATGA	113
Db	531	CTGGTTTCGTCTCTAATCGAGCTGAACACTAGTCACTGGTTCCAGGGTTCTCTTCCGTGA	590

92374	ATCGCTTAAGACACAGCGGGAGGACCAATGATCGGGATATAAACCAGGCAITCCAGCC	923115
789	AGCAACAGCAACCCCTTTGGGTCCTCCCACTGATATGAGGAGCTCTGTTTCACTCTAT	848
92314	GGCAACGGCTACCCCTTTTGGGTCCCTCCCTTGTATGSGAGCTCTGT	92265
849	TTCACTCTATTAATCATCGCAACTGAC--TCCTCTGGTCCGTGTTTTTATGCTCAAG	906
92264	TTCACTCTATTAATCTTGCAACTGCACATTCTTTTGGTCTGTGTTTGTATAGGCTCGAG	92205
907	CTGAGCTTTTGTTCGGCATCCACCACCTGCTGTGTTTGCACCGTCAAGACCCGCTGCTGAC	966
92204	CTGAGCTTTTCGCTCGCGGTCCACCACCTGCTGTTTGGCACTGTTACAGACCTGTGGCTGAC	92145
967	TTCCATCCCTTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCGAGGTACCCAT	1026
92144	TTCCA-CAGGGTGGATCCAGCAGGGGTGTCTGCTGTACTCTGATCCAGTAGGGGCCAC	92086
1027	TGCACCTCCGATCAGGCTTAAGGCTTGGCACTGTTCTTCGCATGCTAAGTGCCTGGTT	1086
92085	TGCTGTCTCCGATAGGGCTACAGGCTTGCATCGTTCCTACATGGCTAAGTGCCTGGTT	92026
1087	TGTCCTTAATAGACTCAACACT---GGTCACTGGGTTCCATGGTTCCTCTCCATGACCC	1142
92025	CGTCTTAATCGAGCTGAACACTAGTCAGTCACTGGGTTCCAAGTTCCTCCGTGACCC	91966
1143	ACGGCTCTTAATAGACTATAACACTCACCGCATGCCCAAGATTCATCTCTTGGTATC	1202
91965	ACGGCTCTTAATAGACTATAACACTCACCGCATGCCCAAGATTCATCTCTTGGAAATC	91906
1203	TGTGAGGCCAAGAACCCCAAGTTCAGAAAGTGTAGGCTTGCCACCATTTGGGAAGTGACC	1262
91905	CGTGAGGCCAAGAACCCCAAGTTCAGAAAGTGTAGGCTTGCCACCATCTTGGAAAGTGACC	91846
1263	CACCTGCCATTTTGGTAGCGGCCCAACACCATCTTGGAGCTGTGGGAGCAAGGATCCCCC	1322
91845	CGCGGCATTTTGGAGTGGCTTGCCACCATCTTGGAGCTCTGGGAGCAAGGAGCACCC	91786
1323	AGTAACA 1329	
91785	GGTAACA 91779	

```

RESULT 12
US-09-949-016-17420
; Sequence 17420, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFECTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17420
; LENGTH: 84571
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84571)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17420

```

Query Match	42.6%	Score 565.6	DB 4	Length 84571
-------------	-------	-------------	------	--------------

[illegible]

RESULT 13
US-09-949-016-15393
; Sequence 15393, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

/	PRIOR APPLICATION NUMBER:	60/237,768
/	PRIOR FILING DATE:	2000-10-03
/	PRIOR APPLICATION NUMBER:	60/231,498
/	PRIOR FILING DATE:	2000-09-08
/	NUMBER OF SEQ ID NOS:	207012
/	SOFTWARE:	FastSeq for Windows Version 4.0
/	SEQ ID NO 15393	
/	LENGTH:	194790
/	TYPE:	DNA
/	ORGANISM:	Human
/	US-09-949-016-	15393
Query Match 42.3%; Score 562; DB 4; Length 194790;		
Best Local Similarity 82.8%; Pred.No. 3.5e-180; Indels 21; Gaps 6;		
Matches 718; Conservative 0; Mismatches 128;		
Qy	409	TAAATGACATTGAAGTCACCCCTCCCAGAGGAAAATCTCAACTGCACAACCCCTTACTACTC 468
Dd	98368	TGATGACATCAAAGGCATCCCTCTTTGAGGMAATCTCAACTGCACAACCCCTACTATGCCCC 98427
Qy	469	CAATTCACTAGNAGCAGTTAGAGCAGTTGTTCAGCGCACCTCCCCAACAGTACTTGGGTT 528
Dd	98428	CAGTTTCAGCAGGAGCAGTTTAGCAGTTGTTCGCGCAACCCTCCCCAAGCAGCATTTGGTT 98487
Qy	529	TTTCCTGTTGAGCGGTGACTCAGAGACAGGACTAGCTGGATTTCCTTAGCGTGAATAAGA 588
Dd	98488	TTTCCATTGAGGG--GGACTGAGAGACAGGACTAGCTGGATTTCCTTAGCGTGAATAAGA 98546
Qy	589	ATCCCNAGGCTANCT-GGGAGGTGACCGCATTCATCTTTTAACATCGGCGTTGCCACT 647
Dd	98547	ATTCTTAAGCTAGCTGGGGAAGGTGACCACACCCCTCTTTTAACACAGAGCTTGTAACT 98606
Qy	648	TAGCTCACACCCGACCAATC-----AGAGAGCTCACTAAAAATGCTAATCAGGC-AA 697
Dd	98607	CAGCTCACACCCGACCAATCAGGTAGTAAAGAGCTCACTAAATACCAATTAGGCTAA 98666
Qy	698	AAACAGGAGTTAAGCAAT-AGCCAATCATCTATTGCTGAGAGCACAGCGGGAAGACA 756
Dd	98667	AAACAGGAGTTAAGAAAATAATCMAATCATCTATCGCTTGAGAGCACAGGGGAGGACA 98726
Qy	757	AGGNTGGGNATAAATCAGCATTCAGCCAGACAAAGCAACCCCTTTGGGTGCCCT 816
Dd	98727	ATGATCGGGATATAACCCAGGATTTGAGCCAGATCAGGTAAACCTCTTTGGGTGCCCT 98786
Qy	817	CCCATTTGATGGGAGCTCTTTTTCACTATTTCATCTATTAATCATGCAACTGCAC 876
Dd	98787	CACACTGATGGGAGCTCTGTTTCACTCTATTAAATCTTGCAACTGCA-----CAC 98838
Qy	877	TCITCTGGTCCGTGTTTTTATGGCTCAAGCTGAGCTTTGTTGCGCATCCAACACTGCT 936
Dd	98839	TCITCTGGTCCATGTTGTTCCGGCTCAAGCTGAGCTTTGCTCGCGTCCACCACTGCT 98898
Qy	937	GTTTGGCCACCGTCACAGACCCGCTGCTGACTTCCTCCATFCCCTTTGGATCCAAGAGTGTC 996
Dd	98899	GMAATGCCGCCATTGCGAGACCTGCCCCTGACTTCCACCCCTCCGATCCGGCAGAGTGTC 98958
Qy	997	ACTGTGCTCTGATCAGCGAGTACCCATTGCCACTCCGATCAGGCTAAAGGCTTCC 1056
Dd	98959	GCTGCACTCTGATCCAAGCGAGGACCCATTGGCACTCCCGATCAGGCTAAAGGCTTCC 99018
Qy	1057	ATTGTTCTCTGATGGCTTAAGTGCTGGGTTGTTCCTTAATAGAACTGAACACTGGTCACTG 1116
Dd	99019	ATTGTTCTCTGCACAGCTAAGTGCTGGGTTCACTCTAATCAGGCTGAACACTGGTGGTG 99078
Qy	1117	GGTTCCATNGTTCTCTTCCATGACCAACGGCTCTTAATAGAGCTATAACACTCACCGAT 1176
Dd	99079	GGTTCCACGGTTCTCTTCCATGACTCACAGCTTCTAATAGAGCTATAACACTCACCAT 99138
Qy	1177	GGCCCAAGATTCCAATCTCTTGTTGTTATCTGTGAGGCCAAGAACCCAGGTCAGAGAANGTA 1236
Dd	99139	GGCCCAAGGTTCCAATCTGTTGGHAATCCATGAGGCCAAGAACCCAGGTCAGAGAATAAA 99198
Qy	1237	GGCTTGGCCACCAATTTGGGAAGTGCCCC 1263

```

Db      99199  GGCGCCGCCCATCTTGGGATGCGCAC  99225
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
RESULT 14
US-08-979-847B-102
; Sequence 102, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMANT:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUBE, PHILIP
;
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC
; THERAPEUTIC PURPOSES
;
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979 847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
;
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-08-979-847B-102

Query Match      41.0%; Score 545; DB 4; Length 635;
Best Local Similarity 92.4%; Pred. No. 4.3e-176;
Matches 572; Conservative 0; Mismatches 47; Indels 0; Gaps 0

QY      257  CCCTGTATCTTCAACTTCCTTGGTTAAAGTTTGTCTCTTCCAGAATTGAAGCTGTAAGCTA  316
DB      1    CCTGTATCTTTAACTTCCTTGTAAAGTTTGTCTCTTCCAGAAATCAAAACTGTAAGAACTA  60

QY      317  CAAATAGTCTTCAAATGGAAACCCAGATGAGTCCATGACTTAAATCTACCGTGGACCC  376
DB      61  CAAATTTGTTCTCAAATGGAGCACAGATGGAGTCCATGACTTAAGATCCACCGTGGACCC  120

QY      377  CTGACCGGCGCTGTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAG  436
DB      121  CTGACCGGCGCTGTAGCCCATGCTCCGATGTTAATGACATTTGAAGGACACCCCTCCCGAG  180

QY      437  GAAATCTCAACTGCAACACCCCTACTACATCTCCAAATTCAGTAGGAAGCAGTTTAGAGCAGT  496

```


Db 181 GAAATCTCACTGACACACCCCTACTATGCCCCCAATTGAGGGAGAGAGGAGT 240
QY 497 TGTAGCCCAACCTCCCAACAGTACTTTGGGTTTCTGTGTAGAGGGTGGACTGAGAGAC 556
Db 241 CATAGCCCAACCTCCCAACAGCAGTCTGGGTTTCTGTGTAGAGGGGGAGTCTGAGAGAC 300
QY 557 AGGACTAGCTGATTTCTTAGCTGACTAAGATCCCAAGGCTTANCTGGGAAGGTGACC 616
Db 301 AGGACTAGCTGATTTCTTAGCTGACTAAGATCCCAAGGCTTANCTGGGAAGGTGACT 360
QY 617 GCATCATCTTTAAACATGGGCTTGCACACTTTAGCTTCACACCGGACCAATCAGAGAGCTC 676
Db 361 GCATCATCTTTAAACATGGGCTTGCACACTTTAGCTTCACACCGGACCAATCAGAGAGCTC 420
QY 677 ACTAAATCTTAATCAGGCAAAACAGAGGATTAAGCAATAGCAATCATCTATTGCTG 736
Db 421 ACTAAATCTTAATCAGGCAAAACAGAGGATTAAGCAATAGCAATCATCTATTGCTG 480
QY 737 AGAGCACAGGGAGGAGCAAGGATGGGATTAAGCAATAGCAATCATCTATTGCTG 796
Db 481 AGAGCACAGGGAGGAGCAAGGATGGGATTAAGCAATAGCAATCATCTATTGCTG 540
QY 797 CAACCCCTTTGGGCTCCCTCCCAATGATGGAGCTCTGTTTCACTCTATTTCACCT 856
Db 541 CAACCCCTTTGGGCTCCCTCCCTTTGATGGGCTCTGTTTCACTCTATTTCACCT 600
QY 857 ATTAATCATGCAACTGCA 875
Db 601 ATTAATCTTGCAACTGAA 619

RESULT 15
US-09-949-016-17002
; Sequence 17002, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17002
; LENGTH: 140224
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(140224)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17002

Query Match 37.6%; Score 500.2; DB 4; Length 140224;
Best Local Similarity 82.4%; Pred. No. 4.6e-159;
Matches 631; Conservative 0; Mismatches 106; Indels 29; Gaps 4;
QY 535 TTGAGGGGTGGAGTACGAGACAGGAGTCTGATTTCTTAGGCTGACTAAGAAATCCCN 594
Db 16074 TAGCAAGCCCTTATTCAGACAGGAGTCTGATTTCTTAGGCCCACTAAGATCCCT 16133
QY 595 AAGCTTANCTGGGAAGGTGACCGCATCTTTTAAACATGGGCTTGCACCTTAGTCTCA 654
Db 16134 AAGCTTANCTGGGAAGGTGACCGCATCTTTTAAACATGGGCTTGCACCTTAGTCTCA 16193
QY 655 CACCCGACCAATC-----AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGA 705

Db 16194 CACCGCAATCAGGTAGTAAAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGA 16253
QY 706 GGTAAAGCAATAGCCCAATCACTATTGCTGAGAGCAGACGCGGAGGAGCAAGGATTGGG 765
Db 16254 GGGAAAGAAATAGCCCAATCACTATTGCTGAGAGCAGACGCGGAGGAGGAGCAATTTGTCAGG 16313
QY 766 ATATAAATCTCAGGCAATCAAGCCAGCAACAGCAACCCCTTTGGGCTCCCTCCCAATTGTA 825
Db 16314 ATATAAATCTCAGGCAATCAAGCCAGCAACAGCAACCCCTTTGGGCTCCCTCCCTTTGTA 16373
QY 826 TGGGAGCTCTGTTTCACTCTATTAAATCATGCAACTGCACTCTCTCTGTT 885
Db 16374 TGGGAGCTCCGTTTCACTCTATTAAAGCTTCCAACTGCA-----CACTCTCTCTGTT 16425
QY 886 CCGTGTCTTTTATGCTCAAGCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGGCCAC 945
Db 16426 CCGTGTCTTTTATGCTCAAGCTGAGCTTTGTTGCGCATCCACCACTGCTGTTTGGCCAC 16485
QY 946 CQTACAGACCCGCT-----GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGT 994
Db 16486 CATCCGACACCCGCGCTGACCCGCGCTGACTTCCACCTTCCAGATCCAGCAGGCTGT 16545
QY 995 CCATCTGCTCTGATCCAGGAGTACCCATTGCGACTCCCGATCAGGCTAAGGCTTG 1054
Db 16546 CTGCTGGGCTCTGATCCAGGAGTGGCCATTGGCGCTCCCGATTGGGCTAAGGCTGG 16605
QY 1055 CCATCTGCTCTGATCCAGGAGTACCCATTGCGACTCCCGATCAGGCTAAGGCTTG 1114
Db 16606 CCATCTGCTCTGATCCAGGAGTGGCTTCCATCTTAATAGAACTGAACACTGCTGTCAC 16665
QY 1115 TGGGTTCCATGTTCTTCTTCCATGACCCAGCGCTTCTTAATAGAGCTATAACACTCACCGC 1174
Db 16666 TGGGTTCCATGTTCTTCTTCCATGACCCAGCGCTTCTTAATAGAGCTATAACACTCACCGC 16725
QY 1175 ATGGCCCAAGATTCCATCTTGGTATCTGTGAGGCAAGAACCCAGCTAGAGAAAGT 1234
Db 16726 GTGGCTTAAGATTCCATCTTGGAAATCCCTGAGGCAAGAA-CCAGGTCAGAAAAACAA 16784
QY 1235 GAGGCTTGCCCAATTTGGGAAGTGGCCCACTGCGCAATTTGGTAGC 1280
Db 16785 GAGGCTTGCCCAATTTGGGAAGTGGCCCACTGCGCAATTTGGTAGC 16830

RESULT 16
US-09-949-016-13413/c
; Sequence 13413, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13413
; LENGTH: 116652
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(116652)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13413


```
Query Match      36.0%; Score 478.6; DB 4; Length 116652;
Best Local Similarity 82.2%; Pred. No. 1.1e-151;
Matches 671; Conservative 0; Mismatches 102; Indels 43; Gaps 9;

QY 547 ACTGAGACAGGACTAGCTGATTTCTAGGCTGACTAAGAATCCCNAAAGCTTANCT-G 605
Db 67040 AGTGAGACAGGACTAGCTGATTTCTAGGCTCACTAAGAATTCCTAAGCCTAGCTGG 66981

QY 606 GGAAGTGACCGCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAA 665
Db 66980 GGAAGTGACCGCACCCACCTTTAAACAAGGGGCTTGGAACGCAGTTACACCTGACCAA 66921

QY 666 TC-----AGAGAGCTCACTAAATGCTAATCAGCG-AAAAACAGGAGGTAAAGCAA 715
Db 66920 TCAGGTAGTAAGAGGGCTCACTAAATACCAATTAGGCTAAAGCAGGAGGTAAAGAA 66861

QY 716 TAG-CCAATCATCTATTGCGTGAGAGCACAGCGGAAGCAAGATTGGGATATAAA-C 773
Db 66860 TAGTCAAAATAATCTATCTGAGAGCACAGGGGAGGGACAATGATTGGGATATAAAC 66801

QY 774 TCAGGCATTCAAGCCAGCA-ACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGGAGC 832
Db 66800 CAGGCAATTCGAGTGGGAGTGGGCAACCCCTTTGGGTCCCTCCCATTTGATGGGAGC 66741

QY 833 TCTGTTTTCACTCTATTTCACTCTATTAAATCATGAACTG--CACTCTTCTGTGTCGGT 890
Db 66740 TCTGT-----TTTCACTCTATTAAATCTTGCAACTGCACACGCTTCTGTGTCGGT 66691

QY 891 TTTTATATGGGCTCAAGCTGAGCTTTTGTTCGCAATCCAACACTGCTGTTTGCACCGTCA 950
Db 66690 TTTGTATGGGCTCGAGCTGAGCTTTCACTCGCTGTCACCACTGCTGTTTGCACCATCC 66631

QY 951 CAGACCCCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACCTGCTCCCTGAT 1010
Db 66630 CAGACCCACTTCTGACTTCCACCTCCCGATCCGGATCCGGCAGGGTGCACCTGCTCCTGAT 66571

QY 1011 CAGCGAGGTACCCATTCGCATCCCGATCAGGCTAAAGGCTTGCCATTGTCCTGCATG 1070
Db 66570 CCAGCAGGACCCATTCGCTCCGATCAGGCTAGAGGCTCGCCATTGTCCTGCATG 66511

QY 1071 GCTAAGTGCCTGGTTGCTCTAATAGAACTGAACACTGGTGCACTGGGTTTCATGGTTCT 1130
Db 66510 GCTAAGTGCCTGGTTGCTCTAATCAAGCTGAACACTAGTGCCTGGTTTCACCGTTCT 66451

QY 1131 CTTCCATGACCCACCGCTTCTAATAGAGCTATAACACTCAGCGATGCCCAAGATTCCA 1190
Db 66450 CTTCTGTACCCATGGCTTCTAATAGAGCTATAACACTCAGCGATGCCCAAGTTCCA 66391

QY 1191 TTCTTTGGTATCTGTGAGGCCAAGAACCCCGAGGTCAAGAAANGTGAGGCTTGCCACCAT 1250
Db 66390 TTCTTTGGAATCCATGAGGCCAAGAACCCCGAGGTCAAGAAANGTTGCTGCCATC 66331

QY 1251 TGGGAAGTGGCC-----CACTGCCATTTTGGTAGCGGCCCAACCAT 1293
Db 66330 TTGGGAGTGGCGCGCCCATTTTGGGAGTGGCGGCCCATCTTTGGGAGAGCCTTCCACCAT 66271

QY 1294 CTTGGGAGCTGGGAGCAAGGATCCCCAGTAACA 1329
Db 66270 CTTGGGAGCTTAAGAGCAAGATCCCTTGGTAACA 66235

RESULT 17
US-09-949-016-11757/c
; Sequence 11757, Application US/09949016
; Patent No. 6812359
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11757
; LENGTH: 131631
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(131631)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11757

Query Match      35.3%; Score 468.6; DB 4; Length 131631;
Best Local Similarity 83.3%; Pred. No. 3.1e-148;
Matches 645; Conservative 0; Mismatches 102; Indels 27; Gaps 9;

QY 545 GGACTGAGACAGGACTAGCTGATTTCTTAGGCTGACTAAGAATCCCNAAAGCTTANCT 604
Db 29989 GAAGTGAGAGACAGAACTAGATGGATTTCTTAGGCCGAGCTAAGAATTTCTTAAGCCTAGCT 29930

QY 605 -GGGAAGGTGACCCGATCCATCTTTAAACATGGGGCTTGCAAACTTAGCTCACACCCGACC 663
Db 29929 GGGGAAGGTGACCCGACTACCTTTTAAACACAGGGCTTGTAACTCAGCTCACACCCGACC 29870

QY 664 AATC-----AGAGAGCTCACTAAATGCTTAATCAGGC-AAAAACAGAGGATTAAGC 713
Db 29869 AATCAGGTAGTAAGAGGGGCTCACTAAATAACAATTTGGGCTAAAGCAGGAGGTAAAGA 29810

QY 714 AATAG-CCAATCATCTATTGCGTGAGAGCACAGCGGGAAGGACAGGATTCGGATATAAA 772
Db 29809 AATAGTCAAAATCATATATTGCTTGAGAGCACAGAGGAGGAGCAATGATCAGGCTATAAA 29750

QY 773 -CTCAGGCATTCAAGCCA-GCAACAGCAAACCCCTTTGGGTCCTCCCTCCCATTTGTATGGGA 830
Db 29749 CCCAGGCATTTCGAGCCAGGTGTGGGCAACCTCTTTGGGTCCTCTCTCTTTGTATGGGA 29690

QY 831 GCTCTGTTTTCACTCTATATTTCACTCTATTAATCATGCAACTG--CACTCTTCTGGTCCG 888
Db 29689 GCCCTGT-----TTTCACTCTATTAATTTGCAACTGCACACTCTCTCTGTTGTCGG 29640

QY 889 TGTTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCACTCCACACTGCTGTTTGCACCGT 948
Db 29639 TGTTTGTTCTGGCTCGAGCTGAGCTTTTCAATTCACAGTCCACCACTGCTGAATGCCACAT 29580

QY 949 CACAGACCCGCTGCTGACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTG 1008
Db 29579 CACAGACCCACCATTTGACTCCACCTCTCCGATCCAGCAGGGTGTCCACTGACCTTCTG 29520

QY 1009 ATCCAGCAGGATACCAATTCGCACTCCCGATCAGGCTTAAGGCTTGCCATTTGTTCTTCA 1068
Db 29519 ATCCAGCAGGATGCCCATCGCACTCTCTGATCGGGCTAGAGGCTCGCCATTTGTTCTTCA 29460

QY 1069 TGGCTAAGTGCCTGGGTTTGTCTTAATAGAACTCAACACTGGTCACTGGGTTCCATGTT 1128
Db 29459 TGGCTAAGTGCCTCAGGTTTCAGGCTTAATCGAGCTGAACACTGATTTGCTGGGTTCCACGTT 29400

QY 1129 CTCTTCCATGACCCACCGGCTTTCTAATAGAGCTATAACACTACCGCATGGCCCCAAGATTC 1189
Db 29399 CTCTTCCATGACCCACCGGCTTTCTAATAGAGCTATAACACTACCGCATGGCCCCAAGGTT 29340

QY 1189 CATTCCTTGGTATCTGTGAGCCCAAGAACCCCGAGTCAGAGAANGTGGGCTTCCACCA 1248
Db 29339 CATTCCTTGGAACTGTGAGCCCAAGAACCCCGAGTCAGAGAACAAGGCTTCCACCA 29280

QY 1249 T-TTGGGAAGTGGGCCACTGCACTTTTGGTAGCGGCCCAACCATCATCTTTGGGAG 1301
Db 29279 TCTTGGGAGTGGGCCCAACCATCTTTGGGAGTGGGCCCAACCATCATCTTTGGGAG 29226
```

RESULT 18
US-09-949-016-17068/c
; Sequence 17068, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17068
; LENGTH: 89584
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17068

Query Match 35.2%; Score 467.2; DB 4; Length 89584;
Best Local Similarity 81.2%; Pred. No. 7.2e-148;
Matches 639; Conservative 0; Mismatches 86; Indels 62; Gaps 6;
QY 545 GCACTCAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCTTACT 604
DB 66010 GAATTAAGACAGGAGTCTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCTTACT 65951

QY 605 GCGAAGTCAACGCTATCTTAAACATGCGGCTTGCAACTTAGCTCACACCCGACCA 664
DB 65950 GCGAAGTCAACGCTATCTTAAACATGCGGCTTGCAACTTAGCTCACACCCGACCA 65891

QY 665 ATC-----AGAGAGTCTACTAAATGCTAATCAGCGAAGGACAGGATTAAGCAA 715
DB 65890 ATCAGGTAGTAAAGAGAGTCTACTAAATGCTAATTAGGCATAA--AGGAGGTAAAGAAA 65833

QY 716 TAGCCAAATCATCTATTGCTGAGACACAGCGGAGGACAGGATTAAGCAA 775
DB 65832 TAGCCAAATCATCTATCCCTGAGACACAGCGTGAAGGACAAATGATCGGGATATAAACCC 65773

QY 776 AGCATTCAAGCAGCAACAGCAACCCCTTTGGGTCCTCCCTTCTCATTTGATGGAGCTCT 835
DB 65772 AGCATTCAAGCAGCAACAGCAACCCCTTTGGGTCCTCCCTTCTCATTTGATGGAGCTCT 65713

QY 836 GTTTTCACTCTATTTCACTCTATTAATCATGCAACTGCA--CTCTTTCTGGTCCGCTGTTT 893
DB 65712 GT-----TTTCACTCTATTAATCATGCAACTGCACTCTCTTTCTGGTCCGCTGTTT 65663

QY 894 TTTATGGCTCAAGCTGAGCTTTTGTGCGCATCCACACTGCTGTTTGGCCACGCTCACAG 953
DB 65662 GTT-----GCGGTCCACCACTGCTGTTTGGCCACGCTCACAG 65627

QY 954 ACCCGTGTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTGATGATCCA 1013
DB 65626 ACTTACTGTGACTTCCATCTCTCCAGATCTGCGAGGTTGCTGACTGCTGATGATCCA 65567

QY 1014 GCGAGGTACCAATTTGCACTCCCGATCAGGCTAAAGGCTTGCCATTTCTCTCATGGCT 1073
DB 65566 GCAAGGCTCCATTTGCTGCTCCCTGATCGGG-----TATTCTGCAAGCT 65522

QY 1074 AAGTGGCTGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGCTCTCTT 1133
DB 65521 GAGTGGCCGGGTTTCAACCTTAATCAAGCTGAACACTAGTCACTGGGTTCCATGGTCCCTT 65462

QY 1134 CCATGACCCACGGCTCTTAATAGAGCTATAAACAATCAGCGCATGGCCCAAGATTCATTC 1193
DB 65461 CCATGACCCACGGCTCTTAATAGAGCTATAAACAATCAGCTGATGAGATTCATTC 65402

QY 1194 CTTGTATCTGTGAGGCCAAGAACCCAGGTGAGAAANGTGAAGCTTTGCCACCACTTTGG 1253
DB 65401 CTTGGAATCCGTGAGGCCAAGAACCCAGGTGAGAAACACAGAGCTTTGCCACCACTTTG 65342

QY 1254 GAAGTGGCCCACTGCCATTTTGTAGCGGCCCAACCACTCTTTGGAGCTGTGGAGCAA 1313
DB 65341 GAAGCAGCTGCGCACCACTTTTGAAGCAGCCTGCGCACCACTCTTTGGAGCTGTGGAGCAA 65282

QY 1314 CGATCCC 1320
DB 65281 GGAATCTC 65275

RESULT 19
US-09-949-016-17057
; Sequence 17057, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17057
; LENGTH: 154023
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17057

Query Match 33.6%; Score 446; DB 4; Length 154023;
Best Local Similarity 81.3%; Pred. No. 2e-140;
Matches 607; Conservative 0; Mismatches 118; Indels 22; Gaps 7;
QY 549 TCAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAAAGCTTANCT-GGG 607
DB 92234 TGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATTCCTAAGTCTATCTGGGG 92293

QY 608 AAGGTGACCGCATCTCATCTTTAAACATGGGGCTTGCAACTAGCTCACACCCGACCAATC 667
DB 92294 AAGGTGACTGCGCCACCTTTAAACATGGGGCTTGTAACCTAGCTCACACCTGACCAATC 92353

QY 668 -----AGAGAGTCTACTAAATGCTAATCAGGC-AAAAACAGGAGGTAAAGCAATA 717
DB 92354 AGGTAATAAGAGAGAGTCTAATAATACCAATAGGCTAAAAAGCAGGAGTAAAGAAATA 92413

QY 718 GCCAA--TCATCTATTGCTGAGAGCACA-GCGGGAAGGACAAGGATTTGGGATATAAATC 775
DB 92414 GTCAAGTCTATCTGCTGAGAGTACCTGCGGAGGACATGATCGGGATATAAATC 92473

QY 776 AGGCATTTCAAGCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGATGGAGCTCT 835
DB 92474 AGGCATTTGAATCGGAGTGGCAACCCCTTTGAGTCCCTCCCTGTTGATAGGAGCTCT 92533

QY 836 GTTTTCACTCTATTTCACTCTATTAATCATGCAACTGCACTC--TTCTGGTCCGCTGTTT 893
DB 92534 GTTTTCACTCTGT-----TAAATAATCTGCACTGCAACCTTTCTGGTCTGTGTTT 92586

QY 894 TTTATGGCTCAAGCTGAGCTTTTGTTCGCACTCCACACTGCTGTTTGGCCACCGTCACAG 953
DB 92587 GTTCTGGCTCAAGCTGAGGTTTCCCTCGCTTCCACACTGCTGTAATGCGGCATTCGAG 92646

QY 954 ACCCGTGTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCA 1013

Db 92647 ACCCACCATTGACTTCCACCCCTCCAGATCTGGCAGGGTATCCACTGCACTTCTGATCCA 92706
 Qy 1014 GCGAGGTACCAATGGCCATCTCCGATCAGGCTAAGGCTTGCCATTTGTTCTGATGGCT 1073
 Db 92707 GGGAGGCGCCCATTTGCCCTCCCGATCAGGCTAGAGGCTTGTCTACTGTTCTCGGTGGCT 92766
 Qy 1074 AAGTGCTGGGTTGTCTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTT 1133
 Db 92767 AAGTGCTGGGTTGTCTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTT 92826
 Qy 1134 CCATGACCCAGGCTTCTAATAGAGCTATAACACTCACCGATGGCCCAAGATCCCAATTC 1193
 Db 92827 CTGTGACCCATGGCTTCTAATAGACTTATAACACTCACTGATGCGCCCAAGGTTCCATTC 92886
 Qy 1194 CTTGGTATCTGTAGGCGCAAGAACCCAGGTCAGAGANGTGAAGCTTGCACCACTTGG 1253
 Db 92887 CTTGAAACTGTGAAGCCAAAGAACCCAGGTCAGAGANGTGAAGCTTGTGTCATCTTT 92946
 Qy 1254 GAAGTGCCCACTGCCATTTTGGTAGC 1280
 Db 92947 GGACAGCTCCACCATCTTGGCAGC 92973

RESULT 20
 US-09-949-016-17296
 ; Sequence 17296, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17296
 ; LENGTH: 251672
 ; TYPE: DNA
 ; ORGANISM: Human
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(251672)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17296

Query Match 33.5%; Score 445.2; DB 4; Length 251672;
 Best Local Similarity 82.4%; Pred. No. 5.3e-140;
 Matches 599; Conservative 0; Mismatches 106; Indels 22; Gaps 7;
 Qy 548 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCCCNAGGCTTANCT-GG 606
 Db 42623 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCTTAAAGCTAGCTGG 42682
 Qy 607 GAAGTGACCGATCCATCTTTAAACATGGGCTTGAACCTTAGCTCACACCCGACCAAT 666
 Db 42693 GAAGTGACCGATCCATCTTTAAACATGGGCTTGAACCTTAGCTCACACCCGACCAAT 42742
 Qy 667 C-----AGAGAGCTCCTAAATGCTTAATCAGGC-AAAAACAGAGGTAAGCAAT 716
 Db 42743 CAGGTAGTAAAGAGGCTCACTAAATACAAATTAGGCTAAGAGCAGGAGTAAAGAAAT 42802
 Qy 717 AG-CCAATCATCTATTGGCTCAGAGCAGCGGAGGACAGAGGATTCGGATATAAA-CT 774
 Db 42803 AGTAAATCATCTATCATCTAGAGCAGCGGGGAGGAGCAATGATTCGGATATAAAACC 42862
 Qy 775 CAGGCATTTCAAGCCAGCA-ACAGCAACCCCTTTGGGTCCCTCCCTCTGATGGAGCT 833

Db 42863 CAGGCATTTCCAGCGGAGTGGCAACCCCTTTGGGTGCCCTCCCTATGATGGAGCT 42922
 Qy 834 CTGTTTTCACCTATTTTCACTCTATTAATCATGCAACTGCACTCTTCTGTCGTTGTTT 893
 Db 42923 CTGTTTTCACCTATTTTCACTCTATTAATCATGCAACTGCACTCTTCTGTCGTTGTTT 42974
 Qy 894 TTTATGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAAG 953
 Db 42975 GTTCTGCTCAAGCTGAGCTTTTCACTCACCATCCACCACTGCTGAATGCCGCCATCAAG 43034
 Qy 954 ACCCGCTGCTCACTTCCATCCCTTTTGGATCCAGCAGAGTGTCCAATGCTCTGATCCA 1013
 Db 43035 ACCCGCTGCTCACTTCTTACCCCTCCAGATCTGGCAGGTTGGGCTGTGCTCTGATCCA 43094
 Qy 1014 GCGAGGTACCAATGGCCATCCCGATCAGGCTAAGGCTTGCCATTTGCTCTGATGGCT 1073
 Db 43095 GCGAGGATCCGTTGCTGCTCTGATTTGGGCTAAGGCTCAACATTTGCTCTGACAGCT 43154
 Qy 1074 AAGTGCTGGGTTTGTCTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTTCTCTT 1133
 Db 43155 AAGTGCTGGGTTTGTCTCTAATAGAACTGAACACTGGTCACTGGGTTCCAGGTTCTCTT 43214
 Qy 1134 CCATGACCCACGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATCCATTC 1193
 Db 43215 CTGTGATCCACAGCTTCTAATAGAGCTATAACACTCACCGCATGGCTCAAGGTTCCATTC 43274
 Qy 1194 CTTGGTATCTGTAGGCGCAAGAACCCAGGTCAGAGANGTGAAGCTTGCACCACTTGG 1253
 Db 43275 CTTGGAATCTGTAGGCGCAAGAACCCCGGCTCAGAGAAACAAAGGCGCCCTCATCTTGG 43334
 Qy 1254 GAAGTGG 1260
 Db 43335 GAGCTGG 43341

Query Match 33.5%; Score 445.2; DB 4; Length 251682;
 Best Local Similarity 82.4%; Pred. No. 5.3e-140;
 Matches 599; Conservative 0; Mismatches 106; Indels 22; Gaps 7;
 Qy 548 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCCCNAGGCTTANCT-GG 606
 Db 42633 CTGAGACAGGACTAGCTGATTTCTTAGCTGACTAGAAATCTTAAAGCTAGCTGG 42692

607 GAAGGTGACCGCATCATCTTTAAACATGCGGCTTGAACCTTAGCTCACACCGACCAAT 666
42693 GAAGGTGACCGCACCCACCTTTAAACACGAGCTTGTAACTCAGCTCACACCAACCAAT 42752
667 C-----AGAGAGCTCAGTAAATGCTAATCAGGC-AAAAACAGAGGTTAAAGCAAT 716
42753 CAGGTAGTAAAGAGGCTCAGTAAATGCTAATCAGGC-AAAAACAGAGGTTAAAGCAAT 42812
717 AG-CCAATCATCTATGCTGAGAGCAGACGCGGAGGAGCAAGGATTCGGATATATAA-CT 774
42813 AGTCAATCATCTATGCTGAGAGCAGACGCGGAGGAGGAGCAAGGATTCGGATATATAA-CT 42872
775 CAGGCAATCAAGCCAGCA-ACAGCAACCCCTTTGGTCCCTCCATGTTGAGGAGCT 833
42873 CAGGCAATCAAGCCAGGAGTGGCAACCCCTTTGGTCCCTCCATGTTGAGGAGCT 42932
834 CTGTTTCACTCTATTTCACTCTATTAATCATGCAACTGCACTCTTTCTGGTCCGTGTTT 893
42933 CTGTTTCACTCTATTAAGCTTGCAACTGCA------CACTCTTCTGGTCTGTGTTT 42984
894 TTTATGGCTCAAGCTGAGCTTTTGTTCGCATCCACCACTGCTGTTTGGCCACCGTCACAG 953
42985 GTTCTGGCTCAAGCTGAGCTTTTCACTCACCATCCACCACTGCTGAATGCCGCATCACAG 43044
954 ACCCGTCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACCTGTGCTCTGATCCA 1013
43045 ACCCGTCTGACTTCTACCCCTTCAGATCTGCGAGGTTGGGCTGCTCTGATCCA 43104
1014 GCGAGGTACCCATGTCACCTCCGATCAGGCTAAAGGCTTGCATGTTCTGCTGATGGCT 1073
43105 GCGAGGCATCCGTTGCTGCTCTGATTTGGCTAAAGGCTCAACATGTTCTGCAAGCT 43164
1074 AAGTGCCTGGTTTGTCTTAATAGAACTGAACTGGTCACTGGGTTCCATGGTCTCTT 1133
43165 AAGTGCCTGGTTTGTCTTAATAGAACTGAACTGGTCACTGGGTTCCATGGTCTCTT 43224
1134 CCATGACCCACGCTTCTAATAGAGCTATAAAGCTTCAAGCTGAGGCTGAGGCTTCCATTC 1193
43225 CTGTATACACAGCTTCTAATAGAGCTATAAAGCTTCAAGCTGAGGCTGAGGCTTCCATTC 43284
1194 CTGTATATCTGTGAGGCTCAAGAACCCAGGCTCAGAGAAGTGGGTTGCCACCAATTTGG 1253
43285 CTGTGAATCTGTGAGGCTCAAGAACCCAGGCTCAGAGAAGTGGGTTGCCACCAATTTGG 43344
1254 GAAGTGG 1260
43345 GAGCTGG 43351

RESULT 22
US-09-573-080A-21
; Sequence 21, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 8523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: (1)..(8523)
; OTHER INFORMATION: her17
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A

TITLE: Prototypic sequences for human repetitive DNA
JOURNAL: Journal of Molecular Evolution
VOLUME: 35
ISSUE: 4
PAGES: 286-291
DATE: 1992-10-
DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
DATABASE ENTRY DATE:
DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-21

Query Match 33.1%; Score 440.4; DB 4; Length 8523;
Best Local Similarity 89.4%; Pred. No. 2.2e-139;
Matches 490; Conservative 0; Mismatches 46; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCCAAAGAGAGGGGAACTGTTTATTTT 60
DB 7988 TCAAAATCGAAGAGCTTTAGACTTCTAAACCGCTCAAGAGAGGGGAACTGTTTATTTT 8047
QY 61 AGGGGAAGAATGCTCTTAGTATGTTAATCAATCTGGAATCAATCTGAGAAAAGTTAAAGA 120
DB 8048 AGGGGAAGAATGCTCTTAGTATGTTAATCAATCTGGAATCGTCACTGAGAAAAGTTAAAGA 8107
QY 121 AATTGAGATCGAATATATATGATAGAGCAGAGGACCTTCAAAACACTGCACTGGGCGCT 180
DB 8108 AATTGAGATCGAATATCAACCGTAGAGCAGAGGAGCTTCGAAAACACTGGGCGCT 8167
QY 181 CCTCAGCCAAATGATGCGCTGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
DB 8168 CCTCAGCCAAATGATGCGCTGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 8227
QY 241 TTTACTCTCTTTTGGACCTGATCTTCAACTTCTTGTAAAGTTTGTCTTCTCCAGAAT 300
DB 8228 GCTACTCTCTTTTGGACCTGATCTTAAACCTCTTGTAACTTGTCTTCTCCAGAAT 8287
QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGAAACCCAGAGTGCAGTCCATGACTAA 360
DB 8288 CGAAGCTGTAAACTA-----CAAAATGAGGCCCAAGATGCAGTCCAAAGACTAA 8335
QY 361 AATCTACCGTGGACCCCTGGACCGCTGCTAGACTATGCTCTGATGTTAATGACATTA 420
DB 8336 GATCTACCGCAGACCCCTGGACCGCTGCTAGCCCAAGATGTTAATGACATTA 8395
QY 421 AGTCAACCCCTCCCGAGGAAATCTCAACTGCAACCCCTTACTACTTCACTTCACTAGG 480
DB 8396 AGGCACCCCTCTCTAGGAAATCTCAGTGTGCAACCTCTACTAGCCCAATTCAGCAGG 8455
QY 481 AAGCAGTTAGAGCAGTTGTGACCCCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA 540
DB 8456 AAGCAGTTAGAGCAGTTGTGACCCCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA 8515
QY 541 GGTGGAC 548
DB 8516 TGGGGGAC 8523

RESULT 23
US-09-949-016-17411/c
; Sequence 17411, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

Fri Feb 25 16:26:28 2005

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17411
; LENGTH: 99580
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17411

Query Match      32.8%; Score 435.8; DB 4; Length 99580;
Best Local Similarity 86.1%; Pred. No. 4.7e-137;
Matches 531; Conservative 0; Mismatches 74; Indels 12; Gaps 4;

QY 245 CTCCTCTTTGGACCCCTGTATCTTCAACTTCCTGTTAAAGTTTGTCTTCCAGAAATTGAA 304
DB 6792 CTACATTTCAATCCCTGTATCTTTAACTCTTGTAAAGTTTGTCTTCCAGAAATCAA 6733

QY 305 GCTGTAAAGTACAAATAGTTCTTCAAAATGGAACCCAGAGTGCATGACATCAAAATC 364
DB 6732 GCTGTAAATATCAAAATTTGTTCTTCAAAATGAACCCAGAGTGCATGACATCAAAATC 6673

QY 365 TACCGTGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTTAAATGACATTTGAAGTC 424
DB 6672 TACCGTGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTTAAATGACATTTGAAGTC 6613

QY 425 ACCCTCCCGAGGAATCTCAACTGCACACCCCTACTACACTCCAATTCAGTAGGAGC 484
DB 6612 ACTCTCTCTGTATGAAATCTAAACTGCAGACCCCTACTATGCCCAATACAGAGGAAGC 6553

QY 485 AGTT-AGAGCAGTTCTCAGCCAACTCTCCCAACAGTACTTGGGTTTCTGTTGAGAGGG 543
DB 6552 ACTTAAGAGCAGTCTGCGCCAACTCTCCCAACAGTACTTGGGTTTCTGTTGAGAGGG 6493

QY 544 TGGACTTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTAAGAAATCCNNAAGCCTTANC 603
DB 6492 GGGACTTGAGAGCAGGACTAGCTGGAATTTCTTAGCTGACTAAGAAATCCNNAAGCCTTANC 6433

QY 604 TGGG-AGGTGACCGCATCTTAAACATGCGGCTTGAACCTTACACTTACACCCGAC 662
DB 6432 TGGGAAAGGTGACCGCATCTTAAACATGCGGCTTGAACCTTACACTTACACCCGAC 6373

QY 663 CAATC-----AGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGGAGTAAAGC 713
DB 6372 CAATCAGGTAGTAAGAGGCTCACTAAATGCTTAATCAGGCAAAACAGGAGTAAAGC 6313

QY 714 AATAG-CCAATCATCTATTGCTGAGACACAGCGGGAAGGACAGGATTTGGATATAA 772
DB 6312 AATAGTCAAAATCATATATCGCTGAGACACAGAGGGGGAACAATGATCAGGATATAA 6253

QY 773 CTCAGGCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGC 832
DB 6252 CTCAGGCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGC 6193

QY 833 TCTGTTTTCACCTATT 849
DB 6192 TCTGTTTTCACCTATT 6176

RESULT 24
US-09-949-016-202370/c
; Sequence 202370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202370
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202370

Query Match      32.7%; Score 434.8; DB 4; Length 601;
Best Local Similarity 86.0%; Pred. No. 2.9e-138;
Matches 520; Conservative 1; Mismatches 69; Indels 15; Gaps 3;

QY 690 TCAGGCAAAACAGGAGGTAAAGCAATAGCAATATGCTTATTCCTCTGAGAGCAGCGGG 749
DB 601 TTAGGCAAAACAGGAGGTAAAGCAATAGCAATATGCTTATTCCTCTGAGAGCAGCGGG 546

QY 750 AAGGCAAAAGGATTGGGATATAAACTCAGGCAITCAAGCCAGCAACAGCAACCCCTTTGG 809
DB 545 AGGCAAAAGGATTGGGATATAAACTCAGGCAITCAAGCCAGCAACAGCAACCCCTTTGG 486

QY 810 GTCCCTCTCCCATTTGATGGAGCTCTGTTTCACTTATTTCACTCTTATTAATCATGCA 869
DB 485 GTCCCTCTCCCATTTGATGGAGCTCTGTTTCACTTATTTCACTCTTATTAATCATGCA 426

QY 870 ACTGCACTCTTCTGGTCCGTTTATATGGCTCAAGCTGAGCTTTTGTTCGCCATCCAC 929
DB 425 ACTGCACTCTTCTGGTCCGTTTATATGGCTCAAGCTGAGCTTTTGTTCGCCATCCAC 366

QY 930 CACTGC-----TGTTTGGCCACCGTCACAGACCCGCTGCTGACTTCCATCCCTTG 979
DB 365 CACTGC-----TGTTTGGCCACCGTCACAGACCCGCTGCTGACTTCCATCCCTTG 306

QY 980 GATCCACAGAGTGTCCACTGTGCTCTGATCCAGCAGGTACCCATTGCCACTCCCGAT 1039
DB 305 GATCYGCAAGGTGTGCTGTGCTCTGATCCAGCAGGTACCCATTGCCACTCCCGAT 246

QY 1040 CAGCTAAAGCTTGGCCATTGTTCTCTCATGGCTTAAGTCTGGGTTTGTCTTAATAGAA 1099
DB 245 TGGCTTAAGCTTGGCCATTGTTCTCTCATGGCTTAAGTCTGGGTTTGTCTTAATAGAA 186

QY 1100 CTGAACACTGTACTGGGTTCCATGTTCTTCTCATGACCCAGCGCTTCTAATAGAGC 1159
DB 185 CTGAACACTGTACTGGGTTCCATGTTCTTCTCATGACCCAGCGCTTCTAATAGAGC 126

QY 1160 TATAACACTCACCACATGGCCATAGATTCCATTCTTGGAAATACCGTGAGGCCAAGAACC 1218
DB 125 TATAACACTCACCACATGGCCATAGATTCCATTCTTGGAAATACCGTGAGGCCAAGAACC 66

QY 1219 CCAGGTGAGAGAAAGTGGGCTTGGCCACCAATTTGGGAAGTGGGCCACTGCCATTTGGTA 1278
DB 65 CCAGGTGAGAGAAAGTGGGCTTGGCCACCAATTTGGGAAGTGGGCCACTGCCATTTGGTA 6

QY 1279 GCGGC 1283
DB 5 GTGGC 1

RESULT 25
US-09-949-016-44299/c
; Sequence 44299, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 44299
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44299

Query Match 32.2%; Score 427.8; DB 4; Length 601;
Best Local Similarity 83.9%; Pred. No. 7.3e-136;
Matches 511; Conservative 1; Mismatches 80; Indels 17; Gaps 2;
QY 538 AGAGGTGACCTGACGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAG 597
DB 601 AGAAGGAGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCGACTAAGATTCCTAAG 542
QY 598 CCTANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACAC 657
DB 541 CCTAGCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCATAT 482
QY 658 CCGACCAATC-----AGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGT 708
DB 481 CTGACCAATCAGGTAGTAAGAGAGTTCATTAATGCTTAATCAGGCAAAAACAGAGGT 422
QY 709 AAAGCAATAGCCATCATCTATTGCTGAGCAGCAGCGGAGGAGGATATGATCAGGATA 768
DB 421 AAAGAAATAGCCATCATCTATTGCTGAGCAGCAGCGGAGGAGGATATGATCAGGATA 362
QY 769 TAACTCAGCATTTAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCATTTGATGG 828
DB 361 TAAACCCAGCATTTGAGCTGGCAATGGCTACCTCTTTGGGTCCCTCCCTTTGATGG 302
QY 829 GAGCTCTGTTTCACTCTATTTCATCTATTAAATCATGCAAACTGCACCTTCTGTCGG 888
DB 301 RAGCTCTGTTTCACTCTATTAAATCTTGAACCTGCA-----CACTCTTTGTCGG 250
QY 889 TGTTTTTATGCTCAAGCTGAGCTTTGTTGGCCATCCACCACCTGCTGTTGCCACCGT 948
DB 249 TGTTTGTACGGCTTGACCTGAGCTTTGCTGTTGCTATCCACCACCTGCTGTTGCCACCGT 190
QY 949 CACAGACCGCTGCTGCTGCTTCCATCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTG 1008
DB 189 TGCAACCCACCACTGACTTCCACCTCTCTGGATCCAGCAGTGTGTTCACTGTGCTCTG 130
QY 1009 ATCCAGCGAGGTACCCATGSCCACTCCCGATCAGGCTAAAGGCTTGCCATTTGTTCTGCA 1068
DB 129 ATCCAGCGAGGCGCCGCTGCCACTCTGATGGGCTAAAGGCTTGCCATTTGTTCTGCA 70
QY 1069 TGCTAAGTCCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTT 1128
DB 69 CGGCTAAGTGCCAGGTTTCACTTAATGGAGCTGAACACTAGTCACTGGGTTCCACGGTT 10
QY 1129 CTCTTCCAT 1137
DB 9 CTCTTCCGT 1

RESULT 26
US-09-949-016-15019
; Sequence 15019, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15019
; LENGTH: 57507
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15019

Query Match 31.0%; Score 411.4; DB 4; Length 57507;
Best Local Similarity 79.6%; Pred. No. 7.6e-129;
Matches 592; Conservative 0; Mismatches 124; Indels 28; Gaps 8;
QY 549 TGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGATCCCNAGCCTTANCT-GGG 607
DB 44269 TGAGAGACAGGACTAGCTGGATTTCTTAGGCGGACTAAGATTTGTTAGCCTAGCTGGGG 44328
QY 608 AAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACTTAGCTCACAACCCGCAATC 667
DB 44329 AAGGTGATCGCACCCACCTTTAAACACCGGGCTTGTAACTCAGCTCACAACCTGACCAATC 44388
QY 668 -----AGAGAGCTCACTAAATGCTTAATCAGG-CAAAAACAGGAGGTAAAGCAATA 717
DB 44389 AGGCAGTAAAGAGGGCTCATTTAAATATCCAAATTAGGTTAAAAAGCAGGAGGTAAAGAAATA 44448
QY 718 G-CCAATCATCTATTGCTGAGCAGCAGCGGAGGAGCAAGATTTGGGATATAAACTC- 775
DB 44449 GTCAATCATCTATCATCTGAGAGCAGGAGGAGGACATGATTTGGGATATAAACCC 44508
QY 776 -AGGCATTTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCTCCCACTTGTATGGGAGCTC 834
DB 44509 AGGCATTTGAGCGGGAGTGGGCCACCCCTTTGGTTCCCTCCCACTGATGGAAGCTC 44568
QY 835 TGTTTTCACTCTATTTCATCTATTAAATCATGCAACTG--CACTCTTCTGTCCTGTT 892
DB 44569 TGT-----TTTCACTCTATTAAATCTTGCACACTGCACACTCTTCTGTTCCACGTT 44618
QY 893 TTTTATGCTCAAGCTGAGCTTTGTTGGCCATCCACCACTGCTGTTTTCGCCACCGTCACA 952
DB 44619 TGTTTTGGCTTGAAGTGGAGCTTTGCTTCCATCTTCACTGCTGTTTACCGGCATCACA 44678
QY 953 GACCGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCC 1012
DB 44679 GACCACTGTTGACTTCCACCCCTCCGGATCCAGCAGGGTGTCTGCTGCAATTTCTGATCC 44738
QY 1013 AGCGAGGTACCAATTGCCACTCCCGATCAGGCTAAAGGCTTGGCAATTTCTTCTGCAATGC 1072
DB 44739 AGCGAGGCGCCCAATTGCTGCT--CATTCGGGCTAGAGGCTGCGCAATTTGTTCTGTGCGGC 44796
QY 1073 TAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGCTGCTGCTGCTGCTGCTGCTCT 1132
DB 44797 TCAGTGTCTGAGTTTGTCTTAATCGAGCTAAACACTAGTTCGGTGGGTTTCATGTTTCTCT 44856
QY 1133 TCATGACCCACCGGCTTCTTAATAGAGCTATAACACTCAGCGATGCGCCCAAGATTTCAAT 1192
DB 44857 TCATGACCCACCGGCTTCTTAATAGAGCTATAACACTCAGCTGATGSCCAAGGTTCAAT 44916
QY 1193 CTTTGGTATCTGTGAGGCCAAGAACCCAGGTCAGAGANGTGAGGCTTGCCACCACTTTG 1252
DB 44917 CTTTGGATCCGTGAGGCCAAGAAATGTTAGGTGAGAGAAACAAAGGCTTGTGCTCCATCT 44976
QY 1253 GGAAGTGGCCCACTGCAATTTGG 1276
DB 44977 GGGAGCAGCCACTTACCAACTTGG 45000

RESULT 27
US-09-949-016-13633/c
; Sequence 13633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13633
; LENGTH: 39686
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(39686)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-13633

```

Query Match	30.1%;	Score 399.8;	DB 4;	Length 39686;
Best Local Similarity	78.1%;	Pred. No. 5.7e-125;		
Matches 598;	Conservative 0;	Mismatches 120;	Indels 48;	Gaps 8;
Qy	549	TGAGAGACAGGACTAGCTGGATTTCCCTAGGCTGACTAAGAATCCCNAAAGCCTTANCT -GGG	607	
Db	6761	TGAGAGCAGGACTAGCTGGATTTCCCTAGGCCGACTAAGAATTCCTAAGCCTAGCTGGGG	6702	
Qy	608	AAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCAACTTAGCTGACACCCGACCAATC	667	
Db	6701	AAGGTGACTGCAACCCACTTTTAAACACGGGGCTTGTAACCTCAGCTCACCTCGACCAATC	6642	
Qy	668	-----AGAGAGCTCAGCTAAAAATGCTAATCAGG -CAAAAAACAGGAGGTGAAGCAATA	717	
Db	6641	AGGTGGTTAAGAAGGCTCAGTAAATATCAATTAGGTTTAAAGCAGGAGTTAAGAAATA	6582	
Qy	718	-GCCAAATCATCTATTGGCTTGAGACGACAGCGGGGAAGCAAGGATTCGGATATATAA -CTC	775	
Db	6581	CTCAAATCATCTATCATCTGAGACGACAGGGGGAGCGACAATGATTGGGATATAAACCCC	6522	
Qy	776	AGGCATTTCAAGCCAGCAACAGCAACCCCTTTTGGGTCGCCCTCCCATTTGATGGGAGCTCT	835	
Db	6521	AGGCATTTGACCAAGGAGGGCAACCCCTTTTGGATCCCTTACCTTTGATATAGCAGCTCT	6462	
Qy	836	GTTTTCACCTCATTTTCACTCATTTAAATCATGCAACTG - -CACTCTTCTGGTCCGTGTTTT	893	
Db	6461	GT-----TTTTCATTTATTAAATCTTGCAACTGCACTCTTTCTGGTCCATGTTTT	6412	
Qy	894	TTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAAG	953	
Db	6411	GTTCTGGCTCGAGCTGAGCTTTTGCTTACCTTCCACCACCTGCTGTTTCGCCACCTGTCAAG	6352	
Qy	954	ACCCGCTGCTGACHTTCCAT-----	990	
Db	6351	ACCCGCTGTTGCCACCGCTCGCAGACCGCTGTTTGACTTCACACCCCTCGATCTGGCAGG	6292	
Qy	991	GTGTCCACTGTGCTCCTGATCCAGGAGGTACCATTCAGTCCCACTCCCGATCAGGCTAAAGG	1050	
Db	6291	GTGTCCGCTGCGTTTCTGATCCAGCGAGGCAACCATGTCGCCATTCCTGTTTGGGCTAGAGG	6232	
Qy	1051	CTTGCCATTGTTCTCTGCATGGCTAAGTGCCCTGGGTTTTGTCTTAATAGAACTGAACATGG	1110	
Db	6231	CTCGCCATTGTTTCTGTGCGGCCAAGTGCCCCAGGTTGCTCTTAATCGAGCTCAACCGTAG	6172	
Qy	1111	TCACTGGGTTCCATGGTTCTCTTCCATGACCCAGCGGCTTCTTAATAGAGCTATTAACACTCA	1170	
Db	6171	TCGCTGGGTTCCAGGATTTCTCTTCCATGACCCATGGCTTCTTAATAGAGCTATTAACACTCA	6112	
Qy	1171	CCGATGCCCCAAGANTCCATTCTCTGGTATCTGTGAGGGCCAGAAACCCAGGTCAGAGA	1230	

```

Db      6111  CTGCATGCCCAAGGTTCCATTCTTTGGAATCCATGAGGCCAAGAGCCTCAGGTCAGAGA 6052
Qy      1231  ANGFTGAGGCTTGCCACCACTTTGGGAACTGGGCCCACTGCGCAATTTGG 1276
Db      6051  AAAAAGGCTTGGCGCCATCTTTGGAGCGCGCTGCCCATCTTTGGG 6006

RESULT 28
US-09-949-016-11770/c
; Sequence 11770, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11770
; LENGTH: 49487
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ...{49487}
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11770

```

Query Match	30.1%	Score 399.8	DB 4	Length 49487	
Best Local Similarity	78.1%	Pred. No. 6.6e-125			
Matches 598	Conservative	0	Mismatches 120	Indels 48	Gaps 8
QY	549	TGAGAGCAGGACTAGCTGGATTTCTTAGCGTGACTAAGAAATCCCNAAAGCCCTANCT--GGG	607		
DB	16051	TGAGAGGCAGGACTAGCTGGATTTCTTAGCGCGACTAAGAAATTCCTAAGCCTAGCTGGGG	15992		
QY	608	AAGGTGACCGCATCCATCTTTTAAACATATGGGGCTTCGAATCTTAGCTCACACCGGACCAATC	667		
DB	15991	AAGGTGACTGCACCCACCTTTTAAACACAGGGGGCTTGTAACTAGCTTCACTCGACCAATC	15932		
QY	668	-----AGAGAGCTCACTAAATGCTTAATCAGG--CAAAAACAGGAGGTAAAGCAATA	717		
DB	15931	AGTGGTAAAGAGGCTCACTAAATATCAATTAGTTTAAAGCAGGAGGTAAAGAAATA	15872		
QY	718	-GCCAATCATCTATTGCCGTGAGACACAGCGGGAAGGACMAGGATTTGGATATAAA--CTC	775		
DB	15871	CTCAAAATCATCTATCATCTGAGAGCACAGGGGGAGCGCAATGATTTGGATATAAACCCC	15812		
QY	776	AGGCATTTCGAAGCAGACACAGCAACCCCTTTTGGGTCCCTCCCACTTTGTATGGGAGCTCT	835		
DB	15811	AGGCATTTGAGCAAGAGAGGGGCAACCCCTTTTGGATCCCTACCTTGTATAGCAGCTCT	15752		
QY	836	GTTTTCACTCTATTTCACTCTATTTAAATCATGCAACTG--CACTCTTCTGCTCCGTGTTT	893		
DB	15751	GT-----TTTCACITTTATTAATCTTGCACTGCACACTCTTCTGGTCCAATGTTT	15702		
QY	894	TTTATGGCTCAAGCTGAGCTTTTGTTCGCGCATCCACCCTGCTGTTTGGCCACCGTCAAG	953		
DB	15701	GTTCTGGCTCGAGCTGAGCTTTTGTCTACCTCCACCCTGCTGTTTGCACACTGTCAAG	15642		
QY	954	ACCGCTGTGTCATTTCAT-----CCCTTTGGATCCAGACA	990		
DB	15641	ACCGCTGTTCGCGCACCGTTCGACAGCCGCTGTGTGACTTCCACCCTCTGGGCTGGCAGG	15582		


```
Db      234  CGTTGCTCCTGATGGGCTAAGGCTCACCATTGTTCTGACAGCTAAGTGCCTGG 175
Qy      1084  GTTTGTCTTAATAGAACTGAACATGGTFCATGGGTTCCATGGTTCCTTCCATGACCCA 1143
Db      174  GTTGTCTTAATCAAGCTGAACACATAGTGTGGTTCACGGTTCCTTCTGTGATCCA 115
Qy      1144  CGGCTTCTAATAGAGCTATACATCCAGCATGGCCAGCAAGATTCCTTCCTGTTATCT 1203
Db      114  CAGCTTCTAATAGAGCTAATACATCACCGCATGGCTCAAGGTTCCATTCCTTGGAACT 55
Qy      1204  GTGAGGCCAAGAACCCAGGTCAGAGAANGTGAGGCTTGGCCACCATTGGGA 1255
Db      54  GTGAGGCCAAGAACCCCGGTCAGAGAACAAAGGCCACCCTCATCTTGGGA 3

RESULT 31
US-09-949-016-44298/c
; Sequence 44298, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44298
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44298

Query Match      27.8%; Score 369.6; DB 4; Length 601;
Best Local Similarity 83.6%; Pred. No. 7e-116;
Matches 432; Conservative 1; Mismatches 76; Indels 8; Gaps 1;

Qy      764  GGATATAAAGCTCAGGCATTCAGCCAGCAAGCAACCCCTTTGGGTCCTCCCATTTG 823
Db      598  GGATATAAAGCCAGGCATTTGAGCTGGCAATGGCTACCCCTCTTTGGGTCCTCCCTTTG 539
Qy      824  TATGGAGCTCTGTTTTCACCTCTATTTCACCTCTATTAAATCATGCAACTGCACTCTCTG 883
Db      538  TATGGAGCTCTGTTTTCACCTCTATTAAATCTTGCAACTGCA-----CACCTTTTG 487
Qy      884  GTCCGTGTTTTTATGGCTCAAGCTGAGCTTTGTTGGCATTCACCACTGCTGTTTGCC 943
Db      486  GTCCGTGTTTGTACGGCTTGACCTGAGCTTTCGCTTGTATCCACCACCTGCTGTTTGCC 427
Qy      944  ACCGTGACAGACCCGCTGCTGACTTCCATTCCTTTGGATCCAGAGAGTGTCCACTGTC 1003
Db      426  ACCGTGACAGACCCCACTGACTTCCACCTCTCTGGATCCAGAGTGTGTTCACTGTGC 367
Qy      1004  TCCTGATCCAGGAGGTACCCATTCGCCACTCCGATCAGGCTAAAGGCTTGCCATTGTTT 1063
Db      366  TCCTGATCCAGGAGGCGCGCTGCCACTCTGTATCGGGCTAAAGGCTTGCCATTGTTT 307
Qy      1064  CTGCATGGCTAAGTGCCTGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTTCCA 1123
Db      306  CTGCAYGGCTAAGTGCCTGGGTTTCACTCTTAATAGGAGCTGAACACTAGTCACTGGGTTCCA 247
Qy      1124  TGGTCTCTTCCATGACCCAGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAA 1183
Db      246  CGGTTCTCTTCGTGACCCATGGCTTCTTAATAGAGCTATAACACTCACTGCATGGCCCAA 187
Qy      1184  GATTCATTTCTTGTGATCTGTGAGGCCAAGAACCCCAAGGTCAGAGAANGTGAGGCTTGC 1243
```

```
Db      186  GATTCATTTCTTTGGAACCTTTTGAGCCAAAGAACCCAGGTCAGAGAAACAAGAGGCTTGC 127
Qy      1244  CACCATTTGGGAAGTGGCCACCACATGCCATTTTGGTAGC 1280
Db      126  CGCTATCTTGAAGGGGCTTGCCACCACATCTTTGGGAGC 90

RESULT 32
US-09-949-016-146564
; Sequence 146564, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146564
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-146564

Query Match      26.0%; Score 345.4; DB 4; Length 601;
Best Local Similarity 83.7%; Pred. No. 1.4e-107;
Matches 405; Conservative 1; Mismatches 68; Indels 10; Gaps 1;

Qy      796  GCAACCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTTCATCTTAATTCACCTC 855
Db      116  GCTACCCCTCTTTGGTCCCTCCCTTTGATGGAGCTCTGTTTTCATCT-----C 165
Qy      856  TATTAAATCATGCAACTGCACTCTTCTGGTCGGTGTGTTTATATGGCTCAAGCTGAGCTTT 915
Db      166  TATTAAATCTTGCAACTGCACTCTTCTGGTCGGTGTGTTTATACGGCTCGAGCTGAGCTTT 225
Qy      916  TGTTCGCCATCCCACTGCTGTTTGCACGCTCACAGACCCGCTGTGACTTCCATCCC 975
Db      226  CACTCTCCATCCCACTGCTGTTTGGCCATCGCAGGCTGCACTGACTTCCATCCC 285
Qy      976  TTTGGATCCAGCAGAGTGTCCACTGTGCTCCTGATCCAGCGAGGTACCCATTGCCATCCC 1035
Db      286  TCTGATCTAGCAGGKTGTCGTTGTCTCTGATCCAGTGAGACGCCCATTTGCCGATCCC 345
Qy      1036  CGATCAGGCTAAAGGCTTGCCATTGTTCTGTCATGGCTAAGTGGCTGGGTTGTCCTAAT 1095
Db      346  CGACTGGGCTAAAGACTTGGCAATTGTTCTTACGCGGCTAAGTGGCCCGGGTTCATCCCTAAT 405
Qy      1096  AGAAGTCAACACTGCTGCTGAGGTTCCATGTTCTCTTCCATGACCCACGCGTTCCTAATA 1155
Db      406  TGAGCTGAACACTAGTCACTGGGTTCCACGGTTCCTCTGTGACCCGCTGAGCTTCTAATA 465
Qy      1156  GAGCTATAACACTCACCGCATGGCCAAAGATTCATTCCTTTGGTATCTGTGAGGCCAAGA 1215
Db      466  GAGCTATAACACTCACCGCGTGGGCCAAAGATTCATTTATTGGAATCCATCAGAGCCAAGA 525
Qy      1216  ACCCAGGTCAGAGAAAGTGGGCTTGGCCACCATTTGGGAGTGGCCACCTGCCATTTTG 1275
Db      526  ACCCAGGTCAGAGAAACACGAGGCTTGGCATCATCTTAGAAGCAGCCCGCCACCATCTTC 585
Qy      1276  GTAG 1279
Db      586  GGAG 589
```



```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 133739
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-133739

Query Match      25.5%; Score 338.8; DB 4; Length 601;
Best Local Similarity 89.9%; Pred. No. 2.6e-105;
Matches 372; Conservative 1; Mismatches 40; Indels 1; Gaps 1;

QY 545 GGACTGAGACAGAGCTAGCTGGATTCCTTAGGCTGACTAAGAAATCCNAGCCTANCT 604
Db 413 GTAGTGAGACAGAGCTAGCTGGATTCCTTAGGCTGACTAAGAAATCCCTAAGCCTAGCT 354
QY 605 GGGAGGTGACCGCATCCATCTTAAACATGGGGCTTCACTTGGCTCAGACCGGACCA 664
Db 353 GGGAAAGGTGACTTCACTCCACCTTTAAACATGGGGCTTCACTTGGCTCAGACCGGACCA 294
QY 665 ATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCAAATC 724
Db 293 ATCAGAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCCAAATC 234
QY 725 ATCTATTGCTGAGAGCAGCGGGAAGACAGAGTGGGATATAAACTCAGGCAATCA 784
Db 233 ATCTATCCCTGAGAGCAGAGGAGGAGCAATGATCAGGATATAAACCCAGGCAATTC 174
QY 785 AGCCAGCAACAGCAACCCCTTGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 844
Db 173 AGCCAGCAACAGCCTACACTCTTTGGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 114
QY 845 CTATTTCACTCTATTAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTATGGCTCA 904
Db 113 CTATTTCACTCTATTAATCTTGCACCTGCACTCTTCTGGTCCATGTTTGTACGGCTCA 54
QY 905 AGCTAGCTTTGTTGGCATCCACCACTGCTGTTTGGCAGCGTCACAGACCG 958
Db 53 AGCTGAAGTTTCGCTCG-CGTCCACCACTGCTGTTTGGCAGCGTGTACAGACCTG 1

RESULT 36
US-09-949-016-120277
; Sequence 120277, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120277
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-120277

Query Match      25.3%; Score 336; DB 4; Length 601;
Best Local Similarity 86.2%; Pred. No. 2.4e-104;
Matches 394; Conservative 1; Mismatches 57; Indels 5; Gaps 2;

QY 877 TCCTTCGCTCGCTGTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCT 936
Db 16 TCCTTCGCTCGCTGTTTTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCT 75
QY 937 GTTTGCCACCGTCACAGACCGCTGCTGCTTCATCCCTTTGGATCCAGAGAGTGTCC 996
Db 76 GTTTGCCACCTGTTACAGACCTGTGGCTGACTTCCA-CAGGGTGGATCCAGCAGGTTGTCT 134
QY 997 ACTGTGCTCTGATCCAGCGAGGTACCCATTTGCCACTCCCGATCAGGCTAAAGCTTGCC 1056
Db 135 GCTGTACTCTCTGATCCAGCTAGGCGCCCACTGCTGCTCCCGATAGGGCTACAGGTTGCC 194
QY 1057 ATTGTTCTCTGATGGCTAAGTGCTCGGTTTGTCTTAATAGAACTGAACACT----GGTC 1112
Db 195 ATCGTTCTACATGGCTAAGTGCTCGGTTTGTCTTAATCGAGCTGAACACTAGTCACTC 254
QY 1113 ACTGGGTTCCATGTTCTCTTCATGACCCACGGCTTCTTAATAGAGCTATTAACACTCACC 1172
Db 255 ACTGGGTTCCACAGTTCTCTTCGTCACCCACGGCTTCTTAATAGAGTATTAACACTCACC 314
QY 1173 GCATGGCCCAAGATTCCATTTCTTGGTATCTGTGAGCCCAAGAACCCAGCTCAGAGAA 1232
Db 315 GCATGGCCCAAGATTCCATTTCTTGGTATCTGTGAGCCCAAGAACCCAGCTCAGAGAA 374
QY 1233 GTGAGGCTTGGCCACCATTTGGGAAAGTGCCCACTGCCATTTTGTAGCGGCCACCA 1292
Db 375 ATGAGGCTTGGCCACCATTTGGGAAAGTGCCCGCCGCTTTTGGAAAGTGGCTTGCACCA 434
QY 1293 TCTTGGAGCTGTGGGAGCAAGGATCCCGCCAGTAACA 1329
Db 435 TCTTGGAGCTGTGGGAGCAAGGACGACCGCGTAAACA 471

RESULT 37
US-09-949-016-202257
; Sequence 202257, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202257
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202257

Query Match      24.4%; Score 324.6; DB 4; Length 601;
Best Local Similarity 82.0%; Pred. No. 2e-100;
Matches 387; Conservative 1; Mismatches 76; Indels 8; Gaps 1;
```



```

Db      292 CATGCCCAAGATTCCATCTTTGGAACTTTGAGCCCAAGAACCCAGGTCAGAGACA 233
QY      1234 TGAGGCTTGCACCACTTTGGGAAGTGGCCCACTGCCATTTTGGTAGC 1280
Db      232 AGAGGCTTGGCGCTATCTTGAAGGGGCTGCGACCATCTTTGGGAGC 186

RESULT 40
US-09-949-016-44296/c
; Sequence 44296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44296
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-44296

Query Match      23.8%; Score 316.4; DB 4; Length 601;
Best Local Similarity 82.2%; Pred. No. 1.3e-97;
Matches 465; Conservative 1; Mismatches 74; Indels 26; Gaps 8;

QY      424 CACCCCTCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGGAAG 483
Db      600 CACTCCTCTGAGGAATCTCAACTGCGAGCCCTACTATGCCCAATTCAGCAGGAAG 541
QY      484 CAGTT-AGAGCAGTTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCTGTGTGAGAGG 542
Db      540 CAGTTAAGAGCGTTGTCGGCCAACTCCCAACAGCAGTGGGTTTTCTGTGTGAGATG 481
QY      543 GTGAGCTGAGAGACAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATCCCNAGCCTAN 602
Db      480 GGGGACTGAGAGACAGGACTAGCTGGATTTCTAGGCTGACTAAGAAATTCCTAAGCCTAG 421
QY      603 CT-GGGAAGGTGACCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGA 661
Db      420 CTGGGGAAGGTGACCGCACCCACCTTTAAACACAGGGTGTGTAAGTCAAGCTGACCTGA 361
QY      662 CCAATC-----AGAGAGCTCAATAAATGCTAATCAGGC-AAAAACAGGAGGTAAA 711
Db      360 CCAATCGGGTAGTAAAGAGGGCTCACTGAAACACAAATTAGGCTAAAGCAGGAGGTAAR 301
QY      712 GCAATAG-CCATATCTATTGCTTGAGAGACAGCGGGAAGGACAAGGATGGGATATA 770
Db      300 GAAACAGTCAATATCATATATCGCTTGAGAGACACAAG-GGGACAATGATCGGGATATA 242
QY      771 AACTCAGGATTCAGGACAGCAACAGCAACCCCTTTGGGTCCCTCCCTATGATGGGA 830
Db      241 AACTCAGGATTCAGGAGGGAGTGGCAACCCCTTTGGGTCCCTCCCTATGATGGGA 182
QY      831 GCTCTGTTTTCACTCTATTTCATCTATTAAATCATGAACCTGCA--CTCTTCTGGTCCG 888
Db      181 GCTCTGT-----TTTCACTCTATTAAATCTTGAACCTGCAACGCTCTCTGCTGT 132
QY      889 TGTTTTTATGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGCACCGT 948
Db      131 TGTTTTTATGCGCTTGTGAGCTGAGCTTTTGTCTCACTGTCCACACCGCTGTTTGC 72
QY      949 CACAGACCGCTGCTGACTTCCATCC 974
```

```

Db      71 TACAGACCTGCCATTGACTCCCAACC 46

Search completed: February 21, 2005, 15:01:37
Job time : 246.166 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 11:52:50 ; Search time 734.518 Seconds
(without alignments)
10694.149 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatcgagagcttttag.....gcaaggtccccagtaaca 1329

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5384158 seqs, 2955248155 residues

Total number of hits satisfying chosen parameters: 10768316

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	99.8	1329	8	US-08-979-847-108
2	1326	99.8	1329	16	US-10-114-104-108
3	1035.2	77.9	56093	10	US-09-873-367C-81
4	798.2	60.1	21646	13	US-10-087-193-910
5	778.8	58.6	1393	17	US-10-220-120-17
6	775.4	58.3	2030	17	US-10-637-565-18
7	774.2	58.3	7974	14	US-10-198-846-9936
8	766.8	57.7	2074	17	US-10-416-642-4
9	731.8	55.1	2946	13	US-10-114-893-134
10	731.8	55.1	2946	15	US-10-016-249-3
11	730.8	55.0	2930	9	US-09-902-535-1

12	722.6	54.4	7582	17	US-10-632-793-30	Sequence 30, Appl
13	721.2	54.3	1136	17	US-10-632-793-25	Sequence 25, Appl
14	712.6	53.6	2782	17	US-10-632-793-26	Sequence 26, Appl
15	711	53.5	2782	17	US-10-133-036-1	Sequence 1, Appl
16	704.8	53.0	1894	9	US-09-864-761-4444	Sequence 4444, Ap
17	643.4	48.4	161334	13	US-10-087-192-730	Sequence 730, App
18	637.8	48.0	780	10	US-09-854-867-385	Sequence 385, App
19	618	46.5	108773	19	US-10-741-600-17907	Sequence 17907, A
20	601.4	45.3	157090	18	US-10-672-764A-34	Sequence 34, Appl
21	579.8	43.6	180557	13	US-10-003-806-6	Sequence 6, Appl
22	579.8	43.6	180557	13	US-10-003-806-9	Sequence 9, Appl
23	570.2	42.9	161671	15	US-10-037-117-1	Sequence 1, Appl
24	545.4	41.0	134292	17	US-10-240-425-1102	Sequence 1102, Ap
25	545	41.0	635	8	US-08-979-847-102	Sequence 102, App
26	545	41.0	635	16	US-10-114-104-102	Sequence 102, App
27	536.2	40.3	849	17	US-10-220-120-15	Sequence 15, Appl
28	530.6	39.9	366710	18	US-10-719-993-6792	Sequence 6792, Ap
29	513.8	38.7	80032	19	US-10-741-600-17727	Sequence 17727, A
30	512	38.5	17758	17	US-10-264-237-2833	Sequence 2833, Ap
31	502	37.8	326014	9	US-09-731-231A-3	Sequence 3, Appl
32	502	37.8	326014	18	US-10-751-985-3	Sequence 3, Appl
33	478.6	36.0	283351	18	US-10-719-993-7065	Sequence 7065, Ap
34	478.6	36.0	283351	19	US-10-741-600-17995	Sequence 17995, A
35	478.6	36.0	1980090	18	US-10-719-993-6815	Sequence 6815, Ap
36	478.6	36.0	1980090	19	US-10-741-600-17676	Sequence 17676, A
37	468.8	35.3	2052	17	US-10-276-774-678	Sequence 678, App
38	465	35.0	3372	17	US-10-632-793-28	Sequence 28, Appl
39	465	35.0	22436	11	US-09-997-722-148	Sequence 148, App
40	460	34.6	285020	13	US-10-087-192-1666	Sequence 1666, Ap
41	457.8	34.4	2145	13	US-10-027-632-98772	Sequence 98772, A
42	457.8	34.4	2145	17	US-10-027-632-98772	Sequence 98772, A
43	455.8	34.3	1003	17	US-10-637-565-1	Sequence 1, Appl
44	455.8	34.3	2052	17	US-10-637-565-14	Sequence 14, Appl
45	453.6	34.1	559	9	US-09-864-761-7501	Sequence 7501, Ap

ALIGNMENTS

RESULT 1

US-08-979-847-108
; Sequence 108, Application US/08979847
; Publication No. US2003039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC AND THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-979-847-108

Query Match 99.8%; Score 1326; DB 8; Length 1329;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATAATCGAAGAGCTTTAGACTTGTCTAAACCCGCAAAAGAGGGGAACTGTTATTTTT 60
DB 1 TCATAATCGAAGAGCTTTAGACTTGTCTAAACCCGCAAAAGAGGGGAACTGTTATTTTT 60

QY 61 AGGGGAAGAATGCTGTTAGTATGTTAACTGGAATCAATCTGGAATCAATCTGGAAGTTAAAGA 120
DB 61 AGGGGAAGAATGCTGTTAGTATGTTAACTGGAATCAATCTGGAAGTTAAAGA 120

QY 121 AATTTGAGATCGAATATATATGATAGAGAGAGGACCTTCAAAACACTGACCCCTGGGGCCT 180
DB 121 AATTTGAGATCGAATATATATGATAGAGAGAGGACCTTCAAAACACTGACCCCTGGGGCCT 180

QY 181 CCTCAGCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 CCTCAGCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240

QY 241 TTTTACTCCTTTTGGACCTGATCTTCAACTTCTTGTAGTTTGTCTTCTTCCAGAAAT 300
DB 241 TTTTACTCCTTTTGGACCTGATCTTCAACTTCTTGTAGTTTGTCTTCTTCCAGAAAT 300

QY 301 TGAAGCTGTAAGAGCTACAATATAGTTCTTCAATGGAACCCCAAGATGCAATGCACTAA 360
DB 301 TGAAGCTGTAAGAGCTACAATATAGTTCTTCAATGGAACCCCAAGATGCAATGCACTAA 360

QY 361 AATCTACCTGGAACCTCGGACCGGCTGCTAGATGATGCTGATGATGATGATGATGATGATGATGAT 420
DB 361 AATCTACCTGGAACCTCGGACCGGCTGCTAGATGATGCTGATGATGATGATGATGATGATGATGAT 420

QY 421 AGTCACCTCCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
DB 421 AGTCACCTCCCGAGGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGG 480

QY 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTGTTGGTTTCTGTTGAGA 540
DB 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTGTTGGTTTCTGTTGAGA 540

QY 541 GGGTGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGATAGATCCCAAGCCT 600
DB 541 GGGTGACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGATAGATCCCAAGCCT 600

QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTAGCTCACACCG 660
DB 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTAGCTCACACCG 660

QY 661 ACCAATCAGAGACTCACTAAATATGTAATCAGGCAAAACAGGAGGTTAAGCAATAGCC 720
DB 661 ACCAATCAGAGACTCACTAAATATGTAATCAGGCAAAACAGGAGGTTAAGCAATAGCC 720

QY 721 AATCATCTATTGCTGAGAGCAGCGGAGGAGCAAGGATTTGGGATATAAACTCAGGCA 780
DB 721 AATCATCTATTGCTGAGAGCAGCGGAGGAGCAAGGATTTGGGATATAAACTCAGGCA 780

QY 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATATGGGAGCTCTGTTTT 840

DB 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATATGGAGCTCTGTTTT 840
QY 841 CACTCTATTTCACTCTATTAAATCATGCAACTCTCTGCTCCGTTGTTTTATGG 900
DB 841 CACTCTATTTCACTCTATTAAATCATGCAACTCTCTGCTCCGTTGTTTTATGG 900

QY 901 CTCAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAACAGACCGCT 960
DB 901 CTCAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCAACAGACCGCT 960

QY 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACTGTGCTCTCTGATCCAGCGAGGT 1020
DB 961 GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCACTGTGCTCTCTGATCCAGCGAGGT 1020

QY 1021 ACCAATGCACTCCCGATCAGCTAAAGCTTTGCCATTTGTTCTTCATGCTGCAATGAGTCC 1080
DB 1021 ACCAATGCACTCCCGATCAGCTAAAGCTTTGCCATTTGTTCTTCATGCTGCAATGAGTCC 1080

QY 1081 TGGGTTTGTCTTAATAGAACTGAACTGCTCACTGGGTTCCATGTTCTCTTCATGAC 1140
DB 1081 TGGGTTTGTCTTAATAGAACTGAACTGCTCACTGGGTTCCATGTTCTCTTCATGAC 1140

QY 1141 CCACGGCTTCTAATAGAGCTATAAACAACCGCATGCCCCCAAGATTCCATTCTCTGGTA 1200
DB 1141 CCACGGCTTCTAATAGAGCTATAAACAACCGCATGCCCCCAAGATTCCATTCTCTGGTA 1200

QY 1201 TCTGTGAGGCAAGAACCCAGGTCAGAGAANGTGAAGCTTGCCACCATTTGGGAAGTGG 1260
DB 1201 TCTGTGAGGCAAGAACCCAGGTCAGAGAANGTGAAGCTTGCCACCATTTGGGAAGTGG 1260

QY 1261 CCCACTGCCATTTTGGTAGCGGCCCAACCATCTTGGGAGCTGTGGAGCAAGGATCCC 1320
DB 1261 CCCACTGCCATTTTGGTAGCGGCCCAACCATCTTGGGAGCTGTGGAGCAAGGATCCC 1320

QY 1321 CCAGTAACA 1329
DB 1321 CCAGTAACA 1329

RESULT 2
US-10-114-104-108
; Sequence 108, Application US/10114104
; Publication No. US20030198647A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/114,104
; FILING DATE: 03-Apr-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,847
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: BERRIDGE, WILLIAM P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 1329 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-10-114-104-108

Query Match 99.8%; Score 1326; DB 16; Length 1329;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAAATCGAAGAGCTTTAGACTGTCTAAACCGCAAAAGAGGGGAAACCTGTTATTTT 60
Db |
QY 1 TCAAAATCGAAGAGCTTTAGACTGTCTAAACCGCAAAAGAGGGGAAACCTGTTATTTT 60
Db |
QY 61 AGGGAGAGATCGTGTAGTATGTTAACTCAATCTGGAATCACTTACTGAGAAAGTTAAAGA 120
Db |
QY 61 AGGGAGAGATCGTGTAGTATGTTAACTCAATCTGGAATCACTTACTGAGAAAGTTAAAGA 120
Db |
QY 121 AATTTGAGATCGAATATAATGTAGACGAGAGGACCTTCAAAACACTGCAACCTCGGGGCT 180
Db |
QY 121 AATTTGAGATCGAATATAATGTAGACGAGAGGACCTTCAAAACACTGCAACCTCGGGGCT 180
Db |
QY 181 CCTCAGCAATGAGTCCCTGGACTCTCCCTTTCTAGGACCTCTAGCAGCTATAATATT 240
Db |
QY 181 CCTCAGCAATGAGTCCCTGGACTCTCCCTTTCTAGGACCTCTAGCAGCTATAATATT 240
Db |
QY 241 TTTTACTCCTTTTGGACCCGTATCTCAACTCTTGTAACTTTGTTCTCTTCCAGAAAT 300
Db |
QY 241 TTTTACTCCTTTTGGACCCGTATCTCAACTCTTGTAACTTTGTTCTCTTCCAGAAAT 300
Db |
QY 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
Db |
QY 301 TGAAGCTGTAAGCTACAAATAGTCTTCAAAATGGAACCCAGATGAGTCCATGACTAA 360
Db |
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db |
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db |
QY 421 AGTCACCCCTCCGAGGAATCTCACTGACACACCCCTACTACTCCTCAATTCAGTAGG 480
Db |
QY 421 AGTCACCCCTCCGAGGAATCTCACTGACACACCCCTACTACTCCTCAATTCAGTAGG 480
Db |
QY 481 AAGCAGTTAGACAGTTGTGACGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db |
QY 481 AAGCAGTTAGACAGTTGTGACGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db |
QY 541 GGGTGGACTCAGAGACAGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 600
Db |
QY 541 GGGTGGACTCAGAGACAGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCNAAAGCCT 600
Db |
QY 601 ANCTGGGAAGGTGACCCGATCCATCTTTAAACATGAGGCTTGAACCTAGCTACACCCG 660
Db |
QY 601 ANCTGGGAAGGTGACCCGATCCATCTTTAAACATGAGGCTTGAACCTAGCTACACCCG 660
Db |
QY 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGGTTAAGCAATAGCC 720
Db |
QY 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGGTTAAGCAATAGCC 720
Db |

1021 ACCCATTCGCACCTCCCGATCAGGCTAAAGGCTTGGCAATTTCTTCTGTCATGCTAGTCC 1080
1021 ACCCATTCGCACCTCCCGATCAGGCTAAAGGCTTGGCAATTTCTTCTGTCATGCTAGTCC 1080
1081 TGGGTTTGTCTTAATAGACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1081 TGGGTTTGTCTTAATAGACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
1141 CCACGGCTTCTTAATAGACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 CCACGGCTTCTTAATAGACTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1201 TCTGTGAGCCCAAGAACCCAGCTCAGAGAAAGTGGAGGCTTGCACCACTTTGGGAAAGTGG 1260
1201 TCTGTGAGCCCAAGAACCCAGCTCAGAGAAAGTGGAGGCTTGCACCACTTTGGGAAAGTGG 1260
1261 CCCACTGCAATTTTGTAGCGGCCCAACCACTTTGGGAGGCTGTGGAGCAAGGATCCC 1320
1261 CCCACTGCAATTTTGTAGCGGCCCAACCACTTTGGGAGGCTGTGGAGCAAGGATCCC 1320
1321 CCAGTAACA 1329
1321 CCAGTAACA 1329

RESULT 3
US-09-873-367C-81
; Sequence 81, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2000-09-29
; PRIOR FILING DATE: 2000-09-29
; PRIOR FILING DATE: 2000-09-29
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 56093

; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-873-367C-81									
Query Match 77.9%; Score 1035.2; DB 10; Length 56093;									
Best Local Similarity 90.1%; Pred. No. 1.6e-310;									
Matches 1133; Conservative 0; Mismatches 111; Indels 13; Gaps 2;									
Qy	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCCAAAGAGAGGGGAACTGTTTATTTTT	60						
Db	37018	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGAACTGTTTATTTTT	37077						
Qy	61	AGGGGAAGATGCTTTAGTAGTATTAATCAATCTGGAATCATTAAGTGAAGTTAAAGA	120						
Db	37078	AGGGGAAGATGCTTTAGTAGTATTAATCAATCTGGAATCATTAAGTGAAGTTAAAGA	37137						
Qy	121	AATTTGAGATCGAATATTAATGTAGAGCAGAGGACCTTCAAAAACATGTCACCCCTGGGGCT	180						
Db	37138	AATTCGAGATCGAATATTAATGTAGAGCAGAGGAGCTTCGAAACACTGGACCTGGGGCT	37197						
Qy	181	CCTCAGCAATGGATGCTTGGATCTCTCCCTTTTAGGACCTCTAGCAGCTATAATAT	240						
Db	37198	CCTCAGCAATGGATGCTTGGATCTCTCCCTTTTAGGACCTCTAGCAGCTATAATAT	37257						
Qy	241	TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTTGTCTCTCCAGAAT	300						
Db	37258	GCTACTCTCTTTGGACCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTCCAGAAT	37317						
Qy	301	TGAAGCTGTAAGCTTACAATAGTTCTTCAAATGGAAACCCAGATGCACTGCATGACTAA	360						
Db	37318	CGAAGCTGTAAGCTTACAATAGTTCTTCAAATGGAAACCCAGATGCACTGCATGACTAA	37365						
Qy	361	AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTGATTTAATGACATTGA	420						
Db	37366	GATCTACCGCAGACCCCTGGACCGGCTGTAGACTATGCTGTGATTTAATGACATTGA	37425						
Qy	421	AGTACCCCTCCGAGGAAGTCTCAATGCAACACCCCTACTACACTCCAATTCAGTAGG	480						
Db	37426	AGGACCCCTCTGAGGAAGTCTCAATGCAACACCCCTACTACACTCCAATTCAGTAGG	37485						
Qy	481	AAGCAGTTAGAGCAGTTGCTCAGCCAACTCCCAACAGTACTTGGGTTTTCTGTGAGA	540						
Db	37486	AAGCAGTTAGAGCAGTTGCTCAGCCAACTCCCAACAGTACTTGGGTTTTCTGTGAGA	37545						
Qy	541	GGGTGGACTGAGAGCAGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT	600						
Db	37546	TGGGGGACTGAGAGCAGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT	37605						
Qy	601	ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCG	660						
Db	37606	AGCTGGGAAGGTGACCATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCTG	37665						
Qy	661	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCMAAAACAGGAGGTAAAGCAATAGCC	720						
Db	37666	ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCMAAAACAGGAGGTAAAGCAATAGCC	37725						
Qy	721	AATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATTTGGATATAAATCTCAGGCA	780						
Db	37726	AATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATTTGGATATAAATCTCAGGTC	37785						
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTT	840						
Db	37786	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTT	37845						
Qy	841	CACCTATTTCACCTATTAAATCATGCAACTGCACTCTTCTGGTCCGTGTTTTTTATGG	900						
Db	37846	CATGCTATTTCACCTATTAAATCTTCAACTGCACTCTTCTGGTCCATGTTTCTTACGG	37905						
Qy	901	CTCAAGCTGAGCTTTTGTTCGGCATTCACCACTAGCTGTTTGGCCCGCTCAGACCCGCT	960						
Db	37906	CTTGAAGCTGAGCTTTTGTTCGGCATTCACCACTAGCTGTTTGGCCCGCTCAGACCCGCT	37965						
Qy	961	GCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCTGCTCTCTGATCCAGCGAGGT	1020						

Db	37966	GCTGACTCCCATCCCTCTGGATCATGCAAGGTGTCCGTGTGCTCCTGATCCAGGAGCC	38025						
Qy	1021	ACCATTTGCCACTCCCGATCAGGCTAAAGGCTTGCCATTGTTCTCTGATGGCTAAGTGCC	1080						
Db	38026	ACCATTTGCCCTCCCAATCGGCTAAAGGCTTGCCATTGTTCTCTGATGGCTAAGTGCC	38085						
Qy	1081	TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTCCATGTTCTCTTCCATGAC	1140						
Db	38086	TGGGTTTGTCTTAATAGAACTGAACACTGGTCACTGGGTCCATGTTCTCTTCTGTCAC	38145						
Qy	1141	CCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCTTGGTA	1200						
Db	38146	CCACGGCTTCTTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTCTT-GAA	38204						
Qy	1201	TCTGTAGGCGCAAGAACCCCGAGTTCAGAAAGTGGAGCTTGCACCAATTTGGGAAG	1257						
Db	38205	TCCATAAGGCGCAAGAACCCCGAGTTCAGAAAGTGGAGCTTGCACCAATTTGGGAG	38261						

RESULT 4

US-10-087-192-910

; Sequence 910, Application US/10087192

; Publication No. US20020182586A1

; GENERAL INFORMATION:

; APPLICANT: Morris, David W.

; APPLICANT: Engelhard, Eric K.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

; FILE REFERENCE: 529452000122

; CURRENT APPLICATION NUMBER: US/10/087,192

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 2059

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 910

; LENGTH: 21646

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)...(21646)

; OTHER INFORMATION: n = A,T,C or G

US-10-087-192-910

Query Match 60.1%; Score 798.2; DB 13; Length 21646;

Best Local Similarity 87.1%; Pred. No. 7.6e-237;

Matches 908; Conservative 0; Mismatches 106; Indels 28; Gaps 2;

Qy	257	CCCTGTATCTTCAACTTCTTGTGTTAAGTTTGTCTTCCAGAAATGAAGCTGTTAAAGCTA	316						
Db	19571	CCCTGTATCTTCAACTTCTTGTGTTAAGTTTGTCTTCCAGAAATGAAGCTGTTAAACTA	19630						
Qy	317	CAATAGTTCTTCAATGGAACCCCGATGCGTCCATGCTAAATCTACCGTGGACCC	376						
Db	19631	CAATAGTTCTTCAATGGAACCCCGATGCGTCCATGCTAAATCTACCGTGGACCC	19690						
Qy	377	CTGGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTTGAAGTCAACCCCTCCCGAG	436						
Db	19691	CTGGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTTGAAGTCAACCCCTCCCGAG	19750						
Qy	437	GAATCTCAACTGCACAAACCCCTACTACCTCCAAATTCAGTAGGAAGCAGTTAGACAGT	496						
Db	19751	GAATCTCAACTGCACAAACCCCTACTATGCCCAATTCAGTAGGAAGCAGTTAGACAGGT	19810						
Qy	497	TGTCAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGAGGGTGGACTGAGAGAC	556						
Db	19811	CATCAGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGAGGGGACTGAGAGAC	19870						
Qy	557	AGGACTAGTGGATTTCTCTAGGCTGACTAAGAAATCCCAAGCCTTANCTGGGAAGGTGACC	616						


```

; PRIOR APPLICATION NUMBER: EP 99420041.8
; PRIOR FILING DATE: 1999-02-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 2030
; TYPE: DNA
; ORGANISM: MSRV-1 retrovirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1626)
US-10-637-565-18

Query Match      58.3%; Score 775.4; DB 17; Length 2030;
Best Local Similarity 92.8%; Pred. No. 2.7e-230;
Matches 812; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 1 TCATAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCGCAAAAGAGGGGGAACCTGTTTATTTT 60
Db 1140 TCATAAATCGAAGAGCTTTAGACTTTGCTAAACCGCCGCAAAAGAGGGGGAACCTGTTTATTTT 1199
Qy 61 AGGGGAAGAATGCTGTTAGTATGTTAATCAATCTGGAATCATTAATCTGAGAAAGTTAAAGA 120
Db 1200 AGGAGAAAGACGCTGTTTATGTTAATCAATCCAGAAATGTCACCTGAGAAAGTTAAAGA 1259
Qy 121 AATTTGAGATCGAATATAATGTTAGACGAGGAGCTTCMAAACACTGCACCTCGGGCCT 180
Db 1260 AATTCGAGATCGAATAAATGTTAGACGAGGAGCTTCMAAACACTGCACCTCGGGCCT 1319
Qy 181 CCTCAGCAATGGATGCCCTGGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db 1320 CCTCAGCAATGGATGCCCTGGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1379
Qy 241 TTTACTCTCTTTGGACCCCTGTAATCTTCAACTCTTCTGTTAAAGTTGTTCTCTCCAGAAAT 300
Db 1380 GTTACTCTCTTTGGACCCCTGTAATCTTCAACTCTTCTGTTAAAGTTGTTCTCTCCAGAAAT 1439
Qy 301 TGAAGCTGTAAGAGCTACAAATAGTCTTCAAAATGGAACCCAGATGCACTGCATGACTAA 360
Db 1440 TGAAGCTGTAAGAGCTACAGATGTTCTTCAAAATGGAACCCAGATGCACTGCATGACTAA 1499
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTTGA 420
Db 1500 GATCCACCGTGGACCCCTGGACCGGCTGCTAGCCATGCTCCGATGTTAATGACTTGA 1559
Qy 421 AGTCACCCCTCCGAGAGAAATCTCAACTGCACACCCCTACTACACTCCAATTCAGTAGG 480
Db 1560 AGGCACCCCTCCGAGAGAAATCTCAACTGCACACCCCTACTATGCCCCCAATTCAGCGGG 1619
Qy 481 AAGCAGTTAGAGCAGTGTGTCAGCCCAACTCCCCAACAGTACTTGGGTTTTCTCTTGAGA 540
Db 1620 AAGCAGTTAGAGCGGTGTCATCAGCCCACTCCCCAACAGCAGCTTGGGTTTTCTCTTGAGA 1679
Qy 541 GGGTGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT 600
Db 1680 GGGGGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCCT 1739
Qy 601 ANCTGGGAAGGTGACCCGCTCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG 660
Db 1740 AGCTGGGAAGGTGACTGCATCCACCTCTAAACATGGGGCTTGCAACTTAGCTCACACCCG 1799
Qy 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
Db 1800 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 1859
Qy 721 AATCATCTATTGCTGAGACGACGCGGGAAGCAAGGATTTGGGATATAAATCAGGCA 780
Db 1860 AATCATCTATTGCTGAGACGACGCGGGAAGCAAGGATTCGGGATATAAATCAGGCA 1919
Qy 781 TTCAGGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGATGGAGCTCTGTTTT 840
Db 1920 TTCAGGCGGCAACGGCAACCCCTTTGGGTCCTCCCTTTGATGGGCGCTCTGTTTT 1979

```

RESULT 6

```

US-10-637-565-18
; Sequence 18, Application US/10637565
; Publication No. US2004004381A1
; GENERAL INFORMATION:
; APPLICANT: PERON Herve
; APPLICANT: PERON Herve
; APPLICANT: KOMURIAN-PRADEL, Florence
; TITLE OF INVENTION: THE LTR REGION OF MSRV-1 AND THE PROTEINS IT ENCODES, AND PROBES
; FILE REFERENCE: 110257
; CURRENT APPLICATION NUMBER: US/10/637,565
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US/09/890,340
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/IB00/00159
; PRIOR FILING DATE: 2000-02-15

```


Db	1203	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACGCCCAAAAGAGGGGAACCTGTTTATTTTT	1262
Qy	61	AGGGAAAGAAATGCTGTTAGTAGTTAACTCAATCTGGAATCATTA	120
Db	1263	AGGAGAAGAAATGCTGTTTATTATGTTAACTCAATCAGAAATGTCTAC	1322
Qy	121	AATTTGAGATCGAATATAAATGTAGAGCAGAGGACCTTCAAAACACTCGACCCCTGGGCGCT	180
Db	1323	AATTCGAGATCGAATACAATGTAGAGCAGAGGAGCTTCAAAACACCCGAAACACTGGGGCGCT	1382
Qy	181	CCTCAGCCAAATGGATGCCCTGGAGCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT	240
Db	1383	CCTCAGCCAAATGGATGCCCTGGGTTCTCCCTCTTTAGGACCTCTAGCAGCTCTAATATT	1442
Qy	241	TTTACTCCTCTTTGGACCTGTATCTTCAACTCTCTTTGAATTTGTTCTCTCTCCAGAAT	300
Db	1443	GTTTACTCCTCTTTGGACCTGTATCTTTAACTCCTTTGTTAAAGTTTGTCTCTTCCAGAAT	1502
Qy	301	TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAGTCCATGACTAA	360
Db	1503	TGAAGCTGTAAAGCTACAAATGGTTCTTCAATGGAGCCCGAGATGCAGTCCATGACTAA	1562
Qy	361	AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTTAATGACATTGA	420
Db	1563	AATCTACCGACGACCCCTGGACCGGCTGCTAGCCCATGCTCCGATGTTAATGACATCGA	1622
Qy	421	AGTCACCCCTCCGAGAGAAATCTCAACTGCACACACCTTACTACACTCCAAATTCAGTAGG	480
Db	1623	AGGCACCTCCTCCCAAGAAAATCTCAACTGCACAAACCCCTACTATGCCCCCAATTCAGCAGG	1682
Qy	481	AAGCAGTTTAGAGCAGTTGTGAGCCAAACCTCCCCAAACAGTACTTGGGTTTTCTGTTGAGA	540
Db	1683	AAGCAGTTTAGCGGTGTCAGTCTCAACCTCCCAACAGCATTGGGTTTTCTGTTGAGA	1742
Qy	541	GGGTGGA	600
Db	1743	GGGGGACTGAGAGACAGGACTAGCTGGATTTCCTAGGCGGATTAAGAAATCCCTTAAGCCT	1802
Qy	601	ANCTGGGAAGGTGACCCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCG	660
Db	1803	AGCTGGGAAGGTGACCGGCTCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCCA	1862
Qy	661	ACCAATCAGAGAGCTCACTAAAATGCTAATCAGGCAAAAAACAGGAGTAAAGCAATAGCC	720
Db	1863	ACCAATCAGAGAGCTCACTAAAATGCTAATTAAGCAAAAAACAGGAGTAAAGCAATAGCC	1922
Qy	721	AATCATCTATTGCTGAGAGACAGCGGAAGGACAAGGATGGGATATAAACTCAGGCA	780
Db	1923	AATCATCTATTGCTGAGAGACAGTGGGAGGGAACAAGGATGGCAATATAAACCCAGGCA	1982
Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCATTTGATGGGAGCTCTGTTTT	840
Db	1983	TTGAGCCAGC-ANAGCAACCGCCTTTGGGTCCCTTCCCTTGTATGGGAGCTCTGTTTT	2041
Qy	841	CAC	873
Db	2042	CAC	2074

2595	TGGGGACTCAGAGACAGGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCCTTAAGCCT	Db
601	ANCTGGGAAGGTGACCGCATCTTTAAACATATGGGCTTTGNACTTTAGCTCACACCCG	QY
2655	AGCTGGGAAGGTGACCAATCCACCTTTTAAACACGGGCTTTGNACTTTAGCTCACACTG	Db
661	ACCAATCAGAGAGCTCACTAAAAATGCTAATCAGGCACAAAAACAGGAGTAAAGCAATAGCC	QY
2715	ACCAATCAGAGAGCTCACTAAATGCTAATTAGGCACAAAAACAGGAGTAAAGAAATAGCC	Db
721	AATCATCTATTGCTTGAGACACAGCGGGAGAGACAGGATTCGGATATATAAATCAAGCA	QY
2775	AATCATCTATTGCTTGAGACACAGAGGAGCAATGATCGGGATATATAACCAAGTC	Db

Qy	781	TTCAAGCCAGCAACAGCAACCCCTTTGGTCCCTCCCAATTGATGGGAGCTCTGTTTT	840
Db	2835	TTGAGCCGCAACGGCAACCCCTTTGGTCCCTCCCTTTGATGGGAGCTCTGTTTT	2894
Qy	841	CACCTATTTTCACTCTCTATAAATCATGCAACTGCA	875
Db	2895	CATGCTATTTTCACTCTCTATAAATCTTGCAACTGCA	2929

RESULT 10
 US-10-016-249-3
 ; Sequence 3, Application US/10016249
 ; Publication No. US20030100053A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Mi, Sha
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6006B.AJ172A
 ; CURRENT APPLICATION NUMBER: US/10/016,249
 ; CURRENT FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: US/03/175,928
 ; PRIOR FILING DATE: 1998-10-20
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2946
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-016-249-3

481	Qy	AAGCAGTTAGACAGATTGTTCAGCCCAACCTCCCCACAGTACTTGGGTTTTCTCTGTTTGAGA	540
2535	Db	AAGCAGTTAGACGGGTCTGTGGCCCAACCTCCCCACAGCACTTAGTGTTCCTGTTTGAGA	2594
541	Qy	GGGTGGACTCAGAGACAGGACTAGCTGGATTTCTCTAGGCTGACTAAGAAATCCCNAAAGCCT	600
2595	Db	TGGGGGACTCAGAGACAGGACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCTTAAGCCT	2654
601	Qy	ANCTGGGAAGGTGACCGGCATTCCTTTTAAACAATGGGGCTTGGCAATTAGCTCACACCCG	660
2655	Db	AGCTGGGAAGGTGACCACATCCACCTTTAAACAGGGGCTTGGCAATTAGCTCACACCTG	2714
661	Qy	ACCAATCAGAGAGCTCACTAAAATGCTTAATCAGGCAAAACACAGAGGTAAGCAATAGCC	720
2715	Db	ACCAATCAGAGAGCTCACTAAAATGCTTAATAGGCAAAAACAGAGGTAAGAAATAGCC	2774
721	Qy	AATCATCTATTTCCTGAGAGCACAGCGGGAAGGACCAAGGATTGGGATATAAACTCAGGCA	780
2775	Db	AATCAICTATTTCCTGAGAGCACAGGAGGAGGACAAATGATCGGATATAAAACCAAGTC	2834
781	Qy	TTCAAGCCACGACACAGCAACCCCTTTTGGGTCCCTCCCAATTGATCGGAGCTCTGTTTT	840
2835	Db	TTCAAGCCGCAACAGGCAACCCCTTTTGGGTCCCTCCCTTTGATCGGAGCTCTGTTTT	2894
841	Qy	CACCTATTTCCTCTATTAAATCATGCAACTGCA	875
2895	Db	CATGCTATTTCCTCTATTAAATCTTTGCAACTGCA	2929

```

RESULT 11
US-09-902-535-1
; Sequence 1, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: Mi, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preeclampsia and gestational trophoblast
; TITLE OF INVENTION: disorders
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2930
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (930) ... (2546)
US-09-902-535-1

```

	Query Match	55.0%;	Score 730.8;	DB 9;	Length 2930;
	Best Local Similarity	90.7%;	Pred. No. 2.7e-216;		
	Matches 793;	Conservative 0;	Mismatches 69;	Indels 12;	Gaps 1
QY	1	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACGCCAAAAGAGGGGGAACCTGTTTATTTTT	60		
DB	2069	TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCTGAAAGAGGGGGAACCTGTTTATTTTT	2128		
QY	61	AGGGGAAGAATGCTGTTAGTAGTATTAATCAATCTGGAAATCAATTACTGAGAAAGTTAAAGA	120		
DB	2129	AGGGGAAGAATGCTGTTTATTATGTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA	2188		
QY	121	AATTTGAGATCGAATAATAATGTACAGACAGAGACCTTCAAAACACTGCACCCTCGGGGCT	180		
DB	2189	AATTCGAGATCGAATACACCTAGACAGAGAGCTTCGAAACACTGCACCCTCGGGGCT	2248		
QY	181	CCTCAGCCAATGGATGCCCTGGACTCTCCCTCTTAGGACCTCTAGCAGGTATAATATT	240		

Db 2249 CCTAGCCAAATGATGCCCTGGATTCTCCCTTCTTAGACCTCTAGCAGCTATAATTT 2308
Qy 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTTGTAAAGTTTGTCTTTCAGAAAT 300
Db 2309 GCTACTCTCTTTGGACCTGATCTTAAACCTCTTGTAACTTGTCTCTTCCAGAA 2368
Qy 301 TGAAGCTGTAAGCTTACAATAGTTCTTCAATGGAACCCAGATGCAATCCATGACTAA 360
Db 2369 CGAAGCTGTAACCTA-----CAATGGAGCCCAAGATGCAAGTCCAAGACTAA 2416
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db 2417 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCA 2476
Qy 421 AGTACCCCTCCCGAGGAAATCTCAACTGCAACCCCTACTACACTCAATTCAGTAGG 480
Db 2477 AGGCACCCCTCTCGAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGCAGG 2536
Qy 481 AAGCAGTTAGACAGTTCTGACGCCAACCTCCCAACAGTCTGGGTTTCTCTGTGAGA 540
Db 2537 AAGCAGTTAGACAGGCTGCTCGGCCAACCTCCCAACAGCCTTAGGTTTCTCTGTGAGA 2596
Qy 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCCTTAGGCTGACTAAGAAATCCNAAGCCT 600
Db 2597 TGGGGGACTGAGAGACAGGACTAGCTGGATTTCCTTAGGCTGACTAAGAAATCCCTAAGCCT 2656
Qy 601 ANCTGGGAAGTGACCGCATCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 2657 AGCTGGGAAGTGACCATCATCCACTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2716
Qy 661 ACCAATCAGAGGCTCACTAAATGCTAATCAGGCAAAACAGGAGTAAAGCAATAGCC 720
Db 2717 ACCAATCAGAGGCTCACTAAATGCTAATTAGGCAAAACAGGAGTAAAGCAATAGCC 2776
Qy 721 AATCATCTATTGCTCTGAGACACAGCGGAAGGACAAGGATTGGGATATAAATCAAGCA 780
Db 2777 AATCATCTATTGCTCTGAGACACAGCAGGAGGACAATGATCGGATATAAATCAAGTC 2836
Qy 781 TTCAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATGTATGGAGGCTGTGTTTT 840
Db 2837 TTCAGCGGCAACCGCAACCCCTTTGGGTCCCTCCCTTTGTATGGGAGCTGTGTTTT 2896
Qy 841 CACTCTATTTCACCTATTAAATCATGCAACTGC 874
Db 2897 CATGCTATTTCACCTATTAAATCTTGCAACTGC 2930

RESULT 12

US-10-632-793-30
; Sequence 30, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauclia
; APPLICANT: MALLET, Francois
; APPLICANT: VOTSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; FILE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30
; LENGTH: 7582
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
; NAME/KEY: misc feature
; LOCATION: (198)..(198)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(355)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1331)..(1331)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2213)..(2213)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2398)..(2398)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3787)..(3787)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4115)..(4115)
; OTHER INFORMATION: n = a or g or c or t/u
FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4261)..(4261)
; OTHER INFORMATION: n = a or g or c or t/u
US-10-632-793-30

Query Match 54.4%; Score 722.6; DB 17; Length 7582;
Best Local Similarity 88.5%; Pred. No. 1.6e-213;
Matches 774; Conservative 19; Mismatches 70; Indels 12; Gaps 1;

Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTAAACCGCAAAAGAGGGGAAACCTGTTATTTT 60
Db 6720 TCRAAATCGAAGAGCTTTAGACTTGTAAACCGCTGARAGAGGGGAAACCTGTTATTTT 6779
Qy 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCAATCTGAAAGTTAAAGA 120
Db 6780 AGGGGAAGATGCTGTTATTAATCAATCTCGGAATCTGCTACTGAAAGTTAAAGA 6839
Qy 121 AATTGAGATCGAATAATAATGTAGACGAGAGGACCTTCAAAACACTGCACCTCGGGCCT 180
Db 6840 AATTCAGATCGAATAACAACGTAKAGCAGAGAGCTTCGAAACACTGCACCTCGGGCCT 6899
Qy 181 CCTCAGCAATGGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGCAGTATAATTT 240
Db 6900 CCTCAGCCRAATGGATGCCCTGGATTCTCCCTTTCTTAGGACCTCTAGCAGTATAATTT 6959
Qy 241 TTTACTCTCTTTGGACCCCTGTTCAACTCTCTTAAAGTTTGTCTCTCCAGAA 300
Db 6960 GCTACTCTCTTTGGACCCCTGTTCTTTACTCTCTTGTAACTTGTCTCTCCAGAA 7019
Qy 301 TGAAGCTGTAAGCTCAAAATAGTTCTTCAAAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db 7020 CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATGCAGTCCAAGACTAA 7067
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATGA 420
Db 7068 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCCGATCTGATGTTAATGACATCAA 7127

QY 421 AGTCAACCCCTCCCGAGGAAATCTCAATGTCACAAACCCCTTAATACTCAATTCAGTAGG 480
DB 7128 AGGCACCCCTCTCTGAGGAAATCTCAGCTGTCACAACTCTACTACGCCCAATTCAGCAGG 7187
QY 481 AAGCAGTTAGACAGTCTGTCAGCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA 540
DB 7188 AAGCAGTTAGACAGTCTGTCAGCAACCTCCCAACAGTACTTGGTCTTCTGTTGAGA 7247
QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGCTAAGAAATCCCAAGCCT 600
DB 7248 TGGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGCTAAGAAATCCCAAGCCT 7307
QY 601 ANCTGGGAAGTGAACCGCATCTTTAAACATGCGGGCTTGCACCTAGCTCACACCG 660
DB 7308 AGSTGGGAAGTGAACCGCATCTTTAAACATGCGGGCTTGCACCTAGCTCACACCTG 7367
QY 661 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
DB 7368 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 7427
QY 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAAAGGATTTGGGATATAAACTCAGGCA 780
DB 7428 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAAAGGATTTGGGATATAAACTCAGGCA 7487
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 840
DB 7488 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 7547
QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGCA 875
DB 7548 CATGCTATTTCACCTCTATTAAATCATGCAACTGCA 7582

RESULT 13
US-10-632-793-25
; Sequence 25, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauca
; APPLICANT: MALLET, Françoise
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT FILING DATE: 2003-08-04
; CURRENT APPLICATION NUMBER: US/10/632,793
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 25
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-25
Query Match 54.3%; Score 721.2; DB 17; Length 1136;
Best Local Similarity 90.0%; Pred. No. 1.6e-213;
Matches 787; Conservative 0; Mismatches 75; Indels 12; Gaps 1;
QY 1 TCAAAATCGAAGCTTTAGACTTGTAAACCGCCAAAGAGGGGACCTGTTATTTT.60
DB 254 TCAAAATCGAAGCTTTAGACTTGTAAACCGCCGTAAGAGAGGGGACCTGTTATTTT 313
QY 61 AGGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
DB 314 AGGGGAAGATGCTGTTATTTATGTTAATCAATCGGNATCGTCACTGAGAAAGTTAAAGA 373

QY 121 AATTGAGATCGAATATAATGTAGAGCAGAGGACCTTCAAAACACTGCACCTCGGGCCT 180
DB 374 AATTGAGATCGAATATAATGTAGAGCAGAGGACCTTCAAAACACTGCACCTCGGGCCT 433
QY 181 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 240
DB 434 CCTCAGCAATGGATGCGCTGAGACTCTCCCTTTCTTAGGACCTCTAGCAGCTATAATTT 493
QY 241 TTTACTCTCTTTTGACCTGATCTTCAACTCTTCTTAAAGTTTGTCTTCCAGAAT 300
DB 494 GCTACTCTCTTTTGACCTGATCTTCAACTCTTCTTAAAGTTTGTCTTCCAGAAT 553
QY 301 TGAAGCTGTAAGCTTCAAAATAGTTCTTCAAAATGGAACCCAGATGCACTCAATGCTAA 360
DB 554 CGAAGCTGTAAGCTTCAAAATAGTTCTTCAAAATGGAACCCAGATGCACTCAATGCTAA 601
QY 361 AATCTACCGTGGACCCCTGGAGCCGCTGCTAGACTATGCTCTGATGTTAATGACATTTGA 420
DB 602 GATCTACCGCAGACCCCTGGAGCCGCTGCTAGCCACGATCTGATGTTAATGACATTTGA 661
QY 421 AGTCAACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAATTCAGTAGG 480
DB 662 AGGCAACCCCTCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAATTCAGTAGG 721
QY 481 AAGCAGTTAGAGCAGTGTGTCAGCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA 540
DB 722 AAGCAGTTAGAGCAGTGTGTCAGCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGA 781
QY 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
DB 782 TGGGGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 841
QY 601 ANCTGGGAAGTGAACCGCATCTTTAAACATGCGGGCTTGCACCTTACTCACAACCCG 660
DB 842 AGTGGGAAGTGAACCGCATCTTTAAACATGCGGGCTTGCACCTTACTCACAACCCG 901
QY 661 ACCAATCAGAGCTCACTATAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
DB 902 ACCAATCAGAGCTCACTATAAATGCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 961
QY 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 780
DB 962 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCA 1021
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 840
DB 1022 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCAATGATGGAGCTCTGTTTT 1081
QY 841 CACTCTATTTCACCTCTATTAAATCATGCAACTGCA 874
DB 1082 CATGCTATTTCACCTCTATTAAATCATGCAACTGCA 1115

RESULT 14
US-10-632-793-26
; Sequence 26, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glauca
; APPLICANT: MALLET, Françoise
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT FILING DATE: 2003-08-04
; CURRENT APPLICATION NUMBER: US/10/632,793
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-26

Query Match      53.6%; Score 712.6; DB 17; Length 2782;
Best Local Similarity 89.4%; Pred. No. 1.2e-210;
Matches 782; Conservative 0; Mismatches 81; Indels 12; Gaps 1;

Qy 1 TC AAAATCGAAGAGCTTTAGACTTGTAAACCGCCAAAGAGGGGAACTGTTATTTT 60
Db 1902 TC GAATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGAACTGTTATTTT 1961

Qy 61 AGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCAATTAAGAGAAAGTTAAAGA 120
Db 1962 AGGGAAGAATGCTGTATTAATGTTAATCAATCCGAATCGTCACTGAGAAAGTTGAAGA 2021

Qy 121 AATTGAGATCGAATATATGATAGAGCAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 2022 AATTCCAGATCGAATATACACGTATAGCAGAGGAGCTTCGAAACACTTGGACCTGGGGCT 2081

Qy 181 CCTCAGCAATGGAATGCGCTGACTCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 240
Db 2082 CCTCAGCGATGGAATGCGCTGGAATCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 2141

Qy 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTTGTCTTCCAGAAAT 300
Db 2142 GCTACTCTCTTTGGACCTGTATCTTGAACCTCTCTTGTAACTTGTCTCTTCCAGAAAT 2201

Qy 301 TGAAGCTGTAAGCTACAATAGTCTTCAATGGAACCCAGATGCAAGTCCATGACTAA 360
Db 2202 CGAAGCTGTAAACTA-----CNAATGGAGCCCAAGATGCAAGTCCAAGACTAA 2249

Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2250 GATCTACCGGAGACCCCTGGACCGGCTGTAGCCGATCTGATGTTAATGACATCA 2309

Qy 421 AGTCAACCCCTCCGAGGAAATCTCAATGCGCAACCCCTACTACACTCCAATTCAGTAGG 480
Db 2310 AGGCACCCCTCTGAGGAAATCTCAGCTGCAACACTCTACTACGCCCAATTCAGCAGG 2369

Qy 481 AAGCAGTTAGAGCAGTTGTAGCCCAACCTCCCAACAGTACTTGGGTTTCTCTTGAGA 540
Db 2370 AAGCAGTTAGAGCAGTTGTAGCCCAACCTCCCAACAGCAGTACTTGGGTTTCTCTTGAGA 2429

Qy 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCT 600
Db 2430 TGGGGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCT 2489

Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 2490 AGGTGGGAAGGTGACCATCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACACCTG 2549

Qy 661 ACCAATCAGAGGCTCACTAAATCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
Db 2550 ACCAATCAGAGGCTCACTAAATCTAATTAGGCAAAACAGGAGGTAAGCAATAGCC 2609

Qy 721 AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTGGGATATAAATCTCAGGCA 780
Db 2610 AATCATCTATTGCTGAGAGCAGCGGGAAGGACAATGATCGGGATATAAATCTCAGGCA 2669

Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATGTATGGAGGCTCTGTTTT 840
Db 2670 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATGTATGGAGGCTCTGTTTT 2729

Qy 841 CACTCTATTTCATCTCTATTAAATCATGCAACTGCA 875
Db 2730 CATGCTATTTCATCTCTATTAAATCTTGCAACTGCA 2764
```

```
US-10-133-036-1
; Sequence 1, Application US/10133036
; Publication No. US20040054133A1
; GENERAL INFORMATION:
; APPLICANT: Conrad, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133,036
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2782
; TYPE: DNA
; ORGANISM: Human endogenous retrovirus
US-10-133-036-1
```

```
Query Match      53.5%; Score 711; DB 17; Length 2782;
Best Local Similarity 89.3%; Pred. No. 3.9e-210;
Matches 781; Conservative 0; Mismatches 82; Indels 12; Gaps 1;

Qy 1 TC AAAATCGAAGAGCTTTAGACTTGTAAACCGCCAAAGAGGGGAACTGTTATTTT 60
Db 1902 TC GAATCGAAGAGCTTTAGACTTGTAAACCGCTGAGAGGGGAACTGTTATTTT 1961

Qy 61 AGGGAAGAATGCTGTAGTATGTTAATCAATCTGGAATCAATTAAGAGAAAGTTAAAGA 120
Db 1962 AGGGAAGAATGCTGTATTAATGTTAATCAATCCGAATCGTCACTGAGAAAGTTGAAGA 2021

Qy 121 AATTGAGATCGAATATATGATAGAGCAGAGGACCTTCAAAACACTGCACCTGGGGCT 180
Db 2022 AATTCCAGATCGAATATACACGTATAGCAGAGGAGCTTCGAAACACTGGAACCTGGGGCT 2081

Qy 181 CCTCAGCAATGGAATGCGCTGACTCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 240
Db 2082 CCTCAGCGATGGAATGCGCTGGATTTCTCCCTTCTTAGGACTCTAGCAGCTATAATATT 2141

Qy 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTTGTCTCTTCCAGAAAT 300
Db 2142 GCTACTCTCTTTGGACCTGTATCTTGAACCTCTCTTGTAACTTGTCTCTTCCAGAAAT 2201

Qy 301 TGAAGCTGTAAGCTACAATAGTCTTCAATGGAACCCAGATGCAAGTCCATGACTAA 360
Db 2202 CGAAGCTGTAAACTA-----CNAATGGAGCCCAAGATGCAAGTCCAAGACTAA 2249

Qy 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
Db 2250 GATCTACCGAAGACCCCTGGACCGGCTGTAGCCGATCTGATGTTAATGACATCAA 2309

Qy 421 AGTCAACCCCTCCGAGGAAATCTCAATGCGCAACCCCTACTACACTCCAATTCAGTAGG 480
Db 2310 AGGCACCCCTCTGAGGAAATCTCAGCTGCAACACTCTACTACGCCCAATTCAGCAGG 2369

Qy 481 AAGCAGTTAGAGCAGTTGTAGCCCAACCTCCCAACAGTACTTGGGTTTCTCTTGAGA 540
Db 2370 AAGCAGTTAGAGCAGTTGTAGCCCAACCTCCCAACAGCAGTACTTGGGTTTCTCTTGAGA 2429

Qy 541 GGGTGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCT 600
Db 2430 TGGGGGACTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCT 2489

Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
Db 2490 AGGTGGGAAGGTGACCATCCACCTTTAAACACCGGGCTTGCAACTTAGCTCACACCTG 2549

Qy 661 ACCAATCAGAGGCTCACTAAATCTAATCAGGCAAAACAGGAGGTAAGCAATAGCC 720
Db 2550 ACCAATCAGAGGCTCACTAAATCTAATTAGGCAAAACAGGAGGTAAGCAATAGCC 2609

Qy 721 AATCATCTATTGCTGAGAGCAGCGGGAAGGACAAGGATTGGGATATAAATCTCAGGCA 780
Db 2610 AATCATCTATTGCTGAGAGCAGCGGGAAGGACAATGATCGGGATATAAATCTCAGGCA 2669

Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATGTATGGAGGCTCTGTTTT 840
Db 2670 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATGTATGGAGGCTCTGTTTT 2729

Qy 841 CACTCTATTTCATCTCTATTAAATCATGCAACTGCA 875
Db 2730 CATGCTATTTCATCTCTATTAAATCTTGCAACTGCA 2764
```

Db 2610 AATCATTTATTCCTGAGGACGACGAGGAGGCAATGATCGGATATAAACCCAAAGTT 2669
Qy 781 TTCAAGCCGCAACAGCAACCCCTTTGGGTCCCTCCCAATGTTATGGAGGCTCTGTTTT 840
Db 2670 TTGAGCGGCAACGCAACCCCTTTGGGTCCCTCCCTTTGTTATGGAGGCTCTGTTTT 2729
Qy 841 CACTTATTTCACTCTTAAATCATGCAACTGCA 875
Db 2730 CATGCTATTTCACTCTTAAATCATGCAACTGCA 2764.

RESULT 16

US-09-864-761-4444
; Sequence 4444, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864, 761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263, 6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4444
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC002346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
US-09-864-761-4444
Query Match 53.0%; Score 704.8; DB 9; Length 1894;
Best Local Similarity 89.7%; Pred. No. 2.7e-208; Mismatches 79; Indels 9; Gaps 1;
Matches 770; Conservative 0;
Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCCAAAGAGGGGAAACCTGTTTATTTT 60
Db 961 TCAAAATCGAAGAGCTTTAGACTTGTCTAACCCGCCAAAGAGGGGAAACCTTTTATTTT 1020
Qy 61 AGGGGAAGAAATGCTGTTAGTATGTTAAATCAATCTGGAATCATCTAGAGAAAGTTAAAGA 120
Db 1021 AGAGGAAAAATGCTGTTGTTATGTTAAATCAATCCGGAATCATCACCGAGAAAGTTAAAGA 1080
Qy 121 AATTTGAGATCGAATATATATGATAGAGGAGGACCTTCAAAACACTGCACCTGGGCGCT 180
Db 1081 AATTCAAGGTGCAATATATAACGTAGAGAAAGAGCTGCAAAACACTGGGACCTTGGGCGCT 1140
Qy 181 CCTCAGCCAAATGGATGCGCTCGACTCTCCCTCTTCTTAGGACCTCTAGCAGCTATAATATT 240
Db 1141 CCTCAGCCAAATGGATGCGCTCGACTCTCCCTCTTCTTAGGACCTCTAGCAGCTATAATATT 1200
Qy 241 TTTACTCTCTTTGGACCTGATCTTCAACTTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 300
Db 1201 GTTACTCTCTTTGGACCTGATCTTCAACTCTCTTGTAAAGTTTGTCTCTTCCAGAAAT 1260
Qy 301 TGAAGCTGTAAAGCTACAATATGTTTCAAAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db 1261 CGAAGCAGTAAAACTACAATCGTTCTTCAAAATGGAGCCCGAGATGCAGTCCATGACTAA 1320
Qy 361 AATCTACCGTGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAAATGACATTTGA 420
Db 1321 AATCTACCGAGGACCCCTGGACCGGCTCTAGACTATGCTCTGATGTTAAATGACATTTGA 1380
Qy 421 AGTCACCCCTCCGAGGAAATCTCAACTGCAACACCCCTACTACATCTCAGTATGAGG 480
Db 1381 AGGCACCCCTCCGAGGAAATCTCAACTGCAACACCCCTACTACATCTCAGTATGAGG 1440
Qy 481 AAGCAGTTAGAGAGTTGTGACCCCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
Db 1441 AAGCAGTTAGAGAGTTGTGACCCCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 1500
Qy 541 GGTGAGCTGAGAGCAGGACTAGCTGGATTTCTTAGGCTGACTAAAGAAATCCCAAGCCT 600
Db 1501 GGGGGAGCTGAGAGCAGGAGTAAGTATGATTTCTTAGACCACTAAGAAATCCCTAGACT 1560
Qy 601 ANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGTTGCAACTTAGCTACACCCG 660
Db 1561 AGCTGGGAAGGTGACCGCTTCCACCTTTAAACACCGGGCTTGCACCTTAGCTACACCCG 1620
Qy 661 ACCAATC-----AGAGAGCTCAATAAATGCTAAATCAGGCAAAACAGAGGATAA 711
Db 1621 ACCAATCAGATACTAAAGAGAGCTCAATAAATGCTAAATGCTAAATGCTAAATGCTAA 1680
Qy 712 GCAATAGCCAATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATTTGGGATATAA 771
Db 1681 GAATAGCCAATCATCTATTGCTGAGAGCAGCGGGAAGGACAGGATTTGGGATATAA 1740
Qy 772 ACTCAGGCAATTCAGCCAGCAACACCCCTTTGGGTCCCTCCCTCCATTTGATGGGAG 831
Db 1741 ACCCAGGCAATTCAGCCAGCTACAGCTACCTCTTTGGGTCCCTCCCTTTGATGGGAG 1800
Qy 832 CTCTGTTTTCATCTATT 849
Db 1801 CTCTGTTTTCATCTATT 1818

RESULT 17
US-10-087-192-730/c
; Sequence 730, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 730
; LENGTH: 161334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(161334)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-730

Query Match 48.4%; Score 643.4; DB 13; Length 161334;
Best Local Similarity 85.4%; Pred. No. 4.1e-188; Indels 46; Gaps 4;
Matches 781; Conservative 0; Mismatches 88;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTAACCGCCAAAGAGGGGAACCTGTTATTTTT 60
Db 100898 TCAAAATGAGAGCTTTAGACTTCTAACCTCTGAAGAGGGGAAGCTGTTATTTTT 100839

QY 61 AGGGGAAGATGCTGTAG-----TATGTTAATCA 90
Db 100838 AGGGGAAGATGCTGTATTTATTTATTTAGCGGAAGATGTTGTTATTTATTTAATCA 100779

QY 91 ATCTGGAATCATTTAGTGAAGATTTAAGAAATTTGAGATCAATATATAGTAGACAGA 150
Db 100778 ATCTGGAATTTACAGAGAAAGTTGAGAAATTCGAGATTTGAATACAGTAGACAGA 100719

QY 151 GGACCTTC-AAAACACTGCACCTCGGGCTCCTCAGCCAAATGGATGCCCTGGACTCTCC 209
Db 100718 GGAGCTTCAAAAACACAGACCTCGGGCTCCTCAGCCAAATGGATGCCCTGGATTCCTC 100659

QY 210 CCTTCTAGGACCTCTAGCAGCTAATAATTTTACTCTCTTTTGGACCTGCTATCTCA 269
Db 100658 CCTTCTAGGATCTCTAGCAGCTTAATAATGATCTCTCTTTGGACCTGATCTTTA 100599

QY 270 ACTTCTCTGTAAGTTTGTCTTCCAGAAATTTGAAGCTGTAAGCTCAAAATAGTTCTTC 329
Db 100598 ACTTCTCTGTAAGTTTGTCTTCCAGAAATTTGAAGCTGTAAGCTCAAAATCGTTCTTC 100539

QY 330 AAATGGAACCCAGATGCAATGATCAATCAATCTACCGTGGACCCCTGGACCGGCTG 389
Db 100538 AAATGGAACCCAGATGCAATGATCAATCAATCTACCGTGGACCCCTGGACCGGCTTA 100479

QY 390 CTAGATATGCTCTGATGTTATGATGATGATGATGATGATGATGATGATGATGATGATG 449
Db 100478 CTAGCCCATGCTCCTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 100419

QY 450 CACAACCCCTACTACACTCCAATTCAGTAGGAGCAGTTAGAGCAGTTGTCAGCCAACT 509
Db 100418 CACAACCCCTACTATGCCCAATTCAGTAGGAGCAGTTAGAGCAGTTGTCAGCCAACT 100359

QY 510 CCCCACAGTACTGTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGCAGACTAGCTGGA 569
Db 100358 CCCCACAGCAGTCTGGGTTTTCTGTTGAGAGGGGAGTGGAGAGCAGGATTTAGCTGGA 100299

QY 570 TTTCTAGGCTGACTAAGATCCNAAGCTTANCTGGAGGTGACCGCATCCATCTTTA 629
Db 100298 TTTCTAGGCGGACTAAGATCCAAAGCTAGCTGGAGGTGACCAATCCACCTTTA 100239

QY 630 AACATGGGGCTTGCAACTTTAGCTTCACCCGACCAATC-----AGAGAGCTACTA 680
Db 100238 AACATGGGGCTTGCAACTTTAGCTTCACCCGACCAATCAGGTAGTAAAGAGAGCTCACTA 100179

QY 681 AAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGCAATCATCTATTGCTGAGAG 740
Db 100178 AAATGCTAATTAGCAAAAACAGGAGGTAAAAAATAGCAATCATCTATCGCTGAGAG 100119

QY 741 CACAGCGGAAGACAAAGGATTCGGGATATAAACTCAGGCAATTCAGCAAGCAAC 800
Db 100118 CACAGCGGAGAGCAATGATCGGATATAAACCCAGGATTCAGGCGGCAACGGCTAC 100059

QY 801 CCCCTTTGGTCCCTCCCATTTGATGGGAGCTCTGTTTTCATCTTATTTCACTCTATTA 860
Db 100058 CTTCTTTGGTCCCTCCCTTTGATGGGAGCTCTCTCT-----GTCTTCACTCTATTA 100005

QY 861 AATCATGCAACTGCA 875
Db 100004 AATATTGCAACTGCA 99990

RESULT 18
US-09-854-867-385
; Sequence 385, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL H
; APPLICANT: ROGAN, PETER K
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/854,867
; CURRENT FILING DATE: 2003-05-08
; NUMBER OF SEQ ID NOS: 613
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 385
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(780)
; OTHER INFORMATION: ltr17
US-09-854-867-385

Query Match 48.0%; Score 637.8; DB 10; Length 780;
Best Local Similarity 90.0%; Pred. No. 1.3e-187; Indels 19; Gaps 2;
Matches 711; Conservative 0; Mismatches 60;

QY 549 TGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCNAAAGCCTTANCTGGGA 608
Db 1 TGAGAGACAGACTAGCTGGATTTCTTAGGCGGACTAAGAATCCCTAGAGCTAGCTGGGA 60

QY 609 AGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTACACCCGACCAATC- 667
Db 61 AGGTGACCGCTTCCACCTTTAAACAGGGGCTTGCAACTTAGCTACACCCGACCAATCA 120

QY 668 -----AGAGAGCTCACTAAATGCTAATCAGGCAAAAACAGGAGGTAAAGCAATAGC 719
Db 121 GGTAGTAAAGAGAGCTCACTAAATGCTAATTAGGCAAAAACAGGAGGTAAAGCAATAGC 180

QY 720 CAATCATCTATTGCTGAGAGCAGACAGCGGGAAGGACAAAGGATTTGGATATAAACTCAGGC 779
Db 181 CAATCATCTATCGCTGAGAGCAGCGGAGGAGGACATGATCGGATATAAAACCCAGGC 240

QY 780 ATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTT 839
Db 241 ATTTCAGCGGCAACGGCTACCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTGT-- 298

QY 840 TCACCTATTTCACCTCTATTAAATCATGCAACTGCACTCTTCTGCTCGGTGTTTTTATG 899

Qy 607 GAAGGTGACCCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCCGACCAAT 666
Db 81135 GAAGGTGACCACTTCCACCTTTAAACACAGGGGCTTGCAACTCAGCTCACACCCGACCAAT 81076
Qy 667 C-----AGAGAGCTCCTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATA 717
Db 81075 CAGATAGTAAGAGAGCTCCTAAATGCTAATTAGGCAAAAGCAGGAGGTAAAGCAATA 81016
Qy 718 GCCAATCATCTATTGCTGAGAGACACAGCGGAAAGGACAAGGATGGGATATAAATCAG 777
Db 81015 GCCAATCATCTATCACCTGAGAGACACAGCGGAGGACAATGATCGGATATAAACCAG 80956
Qy 778 GCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGT 837
Db 80955 GCATTGAGCCGGTAACAGCTACCCCTTTGGGTCCCTCCCTTTGATGGAGCTCTGT 80896
Qy 838 TTTCACCTCTATTTCACCTCTAATAATCATGCAACTGCACCTCTCTGCTCCGTTGTTT 897
Db 80895 -----CTTCACCTCTAATAATCTTGCAACTGCACCTCTCTGCTCCATGTTGTTA 80846
Qy 898 TGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCACCGTCACAGCCC 957
Db 80845 CGGCTCGAGCTTCGCTTTTCGCTCACCGTCCACCACTGCTGTTTGGCCCCGTTGGAGACC 80786
Qy 958 GCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCATCCAGCGA 1017
Db 80785 ACCGCTGGCTTCCATCCCTTCGGATCCCATCAGGGTATCCGCTGCTCTCTGATCCAGCGA 80726
Qy 1018 GGTACCACTTGCACCTCCCGATCAGGCTAAAGGCTTGCATTTGTTCTTCGATGCTAAGT 1077
Db 80725 GCGGCCATTGCCACTCTCGATCGGGCCAAATGCTCGCAATTTGTTCTGCACGGCTAAGT 80666
Qy 1078 GCCTGGGTTTGTCTAATAGAACTGAACACTGGTCACTGGGTTCCATGGTCTCTTCCAT 1137
Db 80665 GCCTGGGTTGCTCTTAATCGAGCTGAACACTAGTCACTGGGTTCCAGGGTTCTCTTCGT 80606
Qy 1138 GACCCACGGCTTCTAATAGAGCTAACAACCTCACCGCATGCCCAAGATTCATTCCTTG 1197
Db 80605 GACCCACGGCTTCTAATAGAGCTAACAACCTCACTGCAATGCGCCCAAGCTTCCATTCCTTG 80546
Qy 1198 GTATCTGTGAGCCCAAGAACCCAGGTGAGAAAGTGAAGCTTGGCCACCATTTGGGAAG 1257
Db 80545 GAATCCGTGAGGCCAAGAACCCAGGTGAGAAACATGAGGCTTGGCCACCATCTTGAAG 80486
Qy 1258 TGGCCCACTGCATTTGGTAGCGCCCAACCATCTTGGAGCTGTTGGAGCAAGGAT 1317
Db 80485 TGGCTGCGGTCATTTTGGAAAGCGGCTTGCCACCATCTTGGAGCTCTGGAGCAAGGC 80426
Qy 1318 CCCCCAGTAACA 1329
Db 80425 CCCCCAATAACA 80414

RESULT 21
US-10-003-806-6/c
; Sequence 6, Application US/10003806
; Publication No. US2002011929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,872
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 180557
; TYPE: DNA
; ORGANISM: Human

US-10-003-806-6

Query Match 43.6%; Score 579.8; DB 13; Length 180557;
Best Local Similarity 86.0%; Pred. No. 2.9e-168;
Matches 683; Conservative 0; Mismatches 90; Indels 21; Gaps 3;
Qy 547 ACTGAGAGACAGACTAGCTGGATTTCTTAGGCTGACTAAGAATCCCAAGACCTTANCTGG 606
Db 58991 AGTGAGAGACAGACTAGCTGGATTTCTTAGGCGGACTAAGATCCCTAAGCTAGCTGG 58932
Qy 607 GAAGGTGACCCGATCCATCTTTAAACATGGGGTTGCAACTTAGCTCACACCCGACCAAT 666
Db 58931 GAAGGTGACCGCTTCCACCTTTAAACATGGGGTTGCAACTTAGCTCACACCCGACCAAT 58872
Qy 667 C-----AGAGAGCTCCTAATAATGCTAATCAGGCAAAACAGAGGTAAAGCAATA 717
Db 58871 CAGATAGTAAGAGAGCTCCTAATAATGCTAATTAGGCAAAACAGAGGTAAAGCAATA 58812
Qy 718 GCCAATCATCTATTGCTGAGAGACACAGCGGAAAGGACAAGGATTTGGGATATAAATCAG 777
Db 58811 GCCAATCATCTATTGCTGAGAGACACAGAGGAGGACAATGATCAGCATATAAACCCAG 58752
Qy 778 GCATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGT 837
Db 58751 GCATCTGAGCCAGCAACAGACTAGCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTGT 58692
Qy 838 TTTCACCTCTATTTCACCTCTAATAATCATCAACTGCA--CTCTTCTGGTCCGTGTTT 895
Db 58691 -----CTTCACCTCTAATAATCTTGAGCTGCACCTCTCTTTGGTCTACATTTGT 58642
Qy 896 TATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACCGTCAAGAC 955
Db 58641 CATGTTTCGAGCTGAGCTTTCTCTCGCGCTCCACCACTGCTGTTTGGCGCTGTCGACAC 58582
Qy 956 CCGTCTGCTGATTCCTCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGC 1015
Db 58581 CTGCTGCTGATTCCTCATCCGTCAGATCCGCAAGTGTCTGCTGTGCTCTCTGATCCAGC 58522
Qy 1016 GAGGTACCCATTCGCACCTCCCGATCAGGCTAAAGGCTTGCCATTTGTTCTCATGGCTAA 1075
Db 58521 GAGCGGCCAATTCAGCTTCCAGTCCGCTTAAGGCTTGCACCTGTTCCACACGGCTAA 58462
Qy 1076 GTGCTCGGGTTGTCCTTAATAGAACTGAACACTCGTCACTGGGTTCCATGTTCTCTTCC 1135
Db 58461 GTGTCGCGGTTGCTCTTAATCAAGCTGAACACTAGTCACTGGGTTCCACGGTCTCTTCC 58402
Qy 1136 ATGACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGCCCAAGATTTCCATTCCT 1195
Db 58401 ATGACCCACGGCTTCTAATAGAGCTATAAACACTCACTGATGGCCCAAGATTCATTCCT 58342
Qy 1196 TGGTATCTGTGAGGCCAAGAACCCCGATCAGAGAAAGTGAAGGCTTGCCACCATTTGGGA 1255
Db 58341 TGGAATCCGTGAGGCCAAGAACCCCGATCAGAGAACACAGAGGCTTGCCACCATCTTGA 58282
Qy 1256 AGTGGCCCATGTCATTTTGGTAGCGGCCCAACCATCTTTGGAGCTGTGGAGCAAGG 1315
Db 58281 AGCAGCCCGCGCATTTTGGAGTGCGGCCCAACCATCTTTGGAGCTCTGGAGCAAGG 58222
Qy 1316 ATCCCCCAGTAACA 1329
Db 58221 ACCACCTGGTAACA 58208

RESULT 22
US-10-003-806-9/c
; Sequence 9, Application US/10003806
; Publication No. US2002011929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066051/10024824

; CURRENT APPLICATION NUMBER: US/10/003,806									
; Sequence 1, Application US/10017117									
; Publication No. US20030124535A1									
; GENERAL INFORMATION:									
; APPLICANT: McCarthy, Jeanette									
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE									
; FILE REFERENCE: MMI-008									
; CURRENT APPLICATION NUMBER: US/10/017,117									
; CURRENT FILING DATE: 2001-12-14									
; PRIOR APPLICATION NUMBER: No. US20030124535A1 yet assigned									
; PRIOR FILING DATE: 2001-12-05									
; NUMBER OF SEQ ID NOS: 3									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 1									
; LENGTH: 161671									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (1)-(161671)									
; OTHER INFORMATION: n = a,t,c or g									
US-10-017-117-1									
Query Match 42.9%; Score 570.2; DB 15; Length 161671;									
Best Local Similarity 88.5%; Pred. No. 2.7e-165;									
Matches 655; Conservative 0; Mismatches 71; Indels 14; Gaps 3;									
QY	547	ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGG	606						
Db	74117	ACTGAGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGG	74176						
QY	607	GAAGGTGACCGCATCCATCTTTAAACATGGGCTTGAACCTAGCTCACACCCGACCAAT	666						
Db	74177	GAAGGTGACCGCATCCATCTTTAAACATGGGCTTGAACCTAGCTCACACCCGACCAAT	74236						
QY	667	CAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCAT	726						
Db	74237	CAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCAT	74296						
QY	727	CTATTGCTGAGAGACAGCGGAGGACAGGATGGGATATAAATCAGGCAATTCAG	786						
Db	74297	CTATTGCTGAGAGACAGCGGAGGACAGGATGGGATATAAATCAGGCAATTCAG	74356						
QY	787	CCAGCAACAGCAACCCCTTTGGGCTCCCTCCCATTTGATGGAGCTCTGTTTCACTCT	846						
Db	74357	CCGGCAATGGCAACCCCTTTGGGCTCCCTCCCATTTGATGGAGCTCTGTTTCACTCT	74407						
QY	847	ATTTCACCTCTATTAAATCAATGCAACTGCACTCTTCTGGTCCGTTGTTTTTATGGCTCAAG	906						
Db	74408	ATTTCACCTCTATTAAATCAATGCAACTGCACTCTTCTGGTCCGTTGTTTTTATGGCTCAAG	74466						
QY	907	CTGAGCTTTTGTGGCATCCACCACTGCTG-TTTGCCACCGTCACAGACCCCGCTGCTGA	965						
Db	74467	CTGAGCTTTTGTGGCATCCACCACTGCTG-TTTGCCACCGTCACAGACCCCGCTGCTGA	74526						
QY	966	CTTCCATCCCTTTGGATCCAGCAGAGTGTCCTGCTGCTGCTGATCCAGGAGGTACCCA	1025						
Db	74527	CTTCCATCCCTTTGGATCCAGCAGAGTGTCCTGCTGCTGCTGATCCAGGAGGTACCCA	74586						
QY	1026	TTGCCACTCCCGATCAGGCTAA---AGGCTTGCCATTGTTCTGCTGATGGCTAAGTGCTG	1082						
Db	74587	TTGCCACTCCCGATCAGGCTAA---AGGCTTGCCATTGTTCTGCTGATGGCTAAGTGCTG	74646						
QY	1083	GGTTTCTCTTAATAGAACTGAACACTGGGTTCATGGTTCATGGTCTCTTCCATGACCC	1142						
Db	74647	GGTTTCTCTTAATAGAACTGAACACTGGGTTCATGGTTCATGGTCTCTTCCATGACCC	74706						
QY	1143	ACGGCTTCTTAATAGAGCTATAACACTCAGGATGCCCAAGATTCCATTCTTCTTGGTATC	1202						
Db	74707	ACGGCTTCTTAATAGAGCTATAACACTCAGGATGCCCAAGATTCCATTCTTCTTGGTATC	74766						
QY	1203	TGTGAGGCCAAGAACCCCGAGGTGAGAGAAAGTGGAGCTTTGCCACCATTTGGGAAGTGGCC	1262						

Db	74767	CGTGAGGCCAAGAACCCCGAGGTCAGAGAACGAGGCTTTGTACCATCTTTGGAAGCAGTC	74826
Qy	1363	CACTGCCATTGTTGGTAGCGG	1282
Db	74827	CACGCCATCTTGGNAGCTG	74846
RESULT 24			
US-10-240-425-1102			
; Sequence 1102, Application US/10240425			
; Publication No. US20040033502A1			
; GENERAL INFORMATION:			
; APPLICANT: Williams, Amanda			
; APPLICANT: Boland, Joseph F.			
; APPLICANT: Lord, Reginald V.			
; APPLICANT: Alvarez, Chris			
; APPLICANT: Wetzel, Jon C.			
; APPLICANT: Scherf, Uwe			
; APPLICANT: Vockley, Joseph G.			
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue			
; FILE REFERENCE: 44921-5026			
; CURRENT APPLICATION NUMBER: US/10/240,425			
; CURRENT FILING DATE: 2002-09-30			
; PRIOR APPLICATION NUMBER: PCT/US01/09847			
; PRIOR FILING DATE: 2001-03-28			
; PRIOR APPLICATION NUMBER: US 60/193,446			
; PRIOR FILING DATE: 2000-03-31			
; NUMBER OF SEQ ID NOS: 1588			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1102			
; LENGTH: 134292			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AL031983			
US-10-240-425-1102			
Query Match 41.0%; Score 545.4; DB 17; Length 134292;			
Best Local Similarity 85.9%; Pred. No. 1.3e-157;			
Matches 680; Conservative 0; Mismatches 89; Indels 23; Gaps 6;			
Qy	547	ACTGAGAGCAGGCTAGCTGATTTCTTAGGCTGACTAAGAAATCCNNAAGCTTANCCTGG	606
Db	54610	AATGAGAGCAGGCTAGCTGATTTCTTAGGCCGCACTAAGAAATCCCTAAGCTAGCTGG	54669
Qy	607	GAAGGTGACCGATCCATCTTTAAACATGGGGCTTGCACTTTAGCTCACACCCGACCAAT	666
Db	54670	GGAGGTGACTGCATCCACCTTTAAACACGGGGCTTGCACTTTAGCTCACACCTGACCAAT	54729
Qy	667	C-----AGAGGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATA	717
Db	54730	CAGGTAGTAAGAGAGCTCACTAAATGCTAATTAGGCAAAACAGGAGGTAAAGATATA	54789
Qy	718	GCAATCATCTATTGCTCAGAGCAGCGGAAGACAAAGATTGGGATATAAACTCAG	777
Db	54790	GCAATCATCTATTGCTCAGATCACAGCGGGAGGACAATGATCGGATATAACACAG	54849
Qy	778	GCATTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTTGATGGAGCTCTGT	837
Db	54850	GCATTCAGCCAGCAAC-GCTACCTCTTTGGGTCCCTTCCCTTTGATGGAGCTCTGT	54908
Qy	838	TTTCACTCTATTTCATCTATTAAATCATGCAACTGCACTCTTCTGGTCGGTGTTTTFA	897
Db	54909	-----CTTCACTCTACTAAATCTTTGCAACTGCACTCTTCTGGTCTATGTTTCTTFA	54958
Qy	898	TGGCTCAAGCTGAGCTTTTGTTCGCCATCCACCATCTGCTGTTTGGCACCCTCACAGACCC	957
Db	54959	CGGCTCGAGGTGAGCTTTGCTTGGCCATCCACCATCTGCGTTTGGCACCCTCGAGACCC	55018
Qy	958	GCTGCTGACTTCATCCCTTTTGGATCCAGCAGAGTGTCACTGTGCTCTCTGATCCAGCA	1017
Db	55019	CGGGCTGACTTCATCCCTTC-GGATCTGGCAGGGTGTCCGCTGTGTCTTCTGAACCAAGTGA	55077

RESULT 25

US-08-979-847-102
; Sequence 102, Application US/08979847
; Publication No. US20030039664A1
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; APPLICANT: BESEME, FREDERIC
; APPLICANT: BEDIN, FREDERIC
; APPLICANT: PARANHOS-BACCALA, GLAUCIA
; APPLICANT: KOMURIAN-PRADEL, FLORENCE
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: MANDRAND, BERNARD
; APPLICANT: GARSON, JEREMY
; APPLICANT: TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
; TITLE OF INVENTION: THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE, PLC
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,847
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

QY 788 CAGCAACAGCAACCCCTTTGGTCCCTCCATGTTATGGAGCTCTGTTTCACTCTA 847
 |||||
 Db 170282 CAGCAATGGCTACCATTTTGGTCCCTCCCTTTGTTATGGAGCTCTGTTTCACTCTA 170223
 |||||
 QY 848 TTTCACTCTATTAAATCATGCAACTGCA 875
 |||||
 Db 170222 TTTCACTCTATTAAATCTTGCACACGCA 170195
 |||||

RESULT 32
 US-10-751-985-3/c
 ; Sequence 3, Application US/10751985
 ; Publication No. US20040126861A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GUEGLER, Karl et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: CL001007CON
 ; CURRENT APPLICATION NUMBER: US/10/751,985
 ; CURRENT FILING DATE: 2004-01-07
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 326014
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(326014)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-751-985-3

Query Match 37.8%; Score 502; DB 18; Length 326014;
 Best Local Similarity 88.7%; Pred. No. 7.3e-144;
 Matches 557; Conservative 0; Mismatches 62; Indels 9; Gaps 1;
 QY 257 CCTGTATCTTCAACTCCCTTGTAAAGTTGCTCTCCAGAAATGAAGCTGTAAGCTTA 316
 Db 170822 CCTGTATCTTCAACTCCCTTGTAAAGTTGCTCTCCAGAAATGAAGCTTCAAAACTA 170763
 QY 317 CAATAGTGTCTTCAAAATGGAACCCAGATGCACTGCACTAAATCTACCTGGACCC 376
 Db 170762 CAATAGTGTCTTCAAAATGGAACCCAGATGCACTGCACTAAATCTACCTGGACCC 170703
 QY 377 CTGACCGGCTGCTAGACTATGCTCTGATGTTAAATGACATGAAGTCACTCCCTCCGAG 436
 Db 170702 CTGACCGGCTGCTAGACTATGCTCTGATGTTAAATGACATGAAGTCACTCCCTCCGAG 170643
 QY 437 GAAATCTCAACTGCACACCCCTACTACACTCCAATTCAGTAGGAGCGATTTAGACAGT 496
 Db 170642 GAAATCTCAACTGCACACCCCTACTACACTCCAATTCAGTAGGAGCGATTTAGACAGT 170583
 QY 497 TGTGACCACTCCCAACAGTACTTGGGTTTCTGTTGAGAGGCTGACTGAGAGAC 556
 Db 170582 CATCGGCATCTTCCCAACAGTACTTGGGTTTCTGTTGAGAGGCTGACTGAGAGAC 170523
 QY 557 AGGACTAGCTGGAATTTCTAGGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGACC 616
 Db 170522 AGGACTAGCTGGAATTTCTAGGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGACT 170463
 QY 617 GCATCATCTTTAAACATGGGGCTTGCACCTTAGCTCACACCCGACCAATC----- 667
 Db 170462 GCTTCTACCTTTAAACCCGGGGCTTGCACCTTAGCTCACACCCGACCAATC----- 170403
 QY 668 AGAGAGCTCACTAAATGCTAAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATATCATC 727
 Db 170402 AGAGAGCTCACTAAATGCTAAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATATCATC 170343
 QY 728 TATTGGCTGAGACACACCGGGAGGACAAAGGATTTGGGATATAAATTCAGGCAATTCAGC 787
 Db 170342 TATCGCTGAGACACACAGTGGGAGGACAAATGTTGGGTTATAAACCCAGGCAATTCAGC 170283

QY 788 CAGCAACAGCAACCCCTTTGGTCCCTCCATGTTATGGAGCTCTGTTTCACTCTA 847
 |||||
 Db 170282 CAGCAATGGCTACCATTTTGGTCCCTCCCTTTGTTATGGAGCTCTGTTTCACTCTA 170223
 |||||
 QY 848 TTTCACTCTATTAAATCATGCAACTGCA 875
 |||||
 Db 170222 TTTCACTCTATTAAATCTTGCACACGCA 170195
 |||||

RESULT 33
 US-10-719-993-7065
 ; Sequence 7065, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001496
 ; CURRENT APPLICATION NUMBER: US/10/719,993
 ; CURRENT FILING DATE: 2003-11-24
 ; NUMBER OF SEQ ID NOS: 55342
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7065
 ; LENGTH: 283351
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(283351)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
 US-10-719-993-7065

Query Match 36.0%; Score 478.6; DB 18; Length 283351;
 Best Local Similarity 90.1%; Pred. No. 1.3e-136;
 Matches 534; Conservative 1; Mismatches 52; Indels 6; Gaps 2;
 QY 283 GTTGTCTCTCCAGAAATGAAGCTGTAAAGCTTAAAGTGTCTTCAATAGTGTCTTCAATGGAACCCCA 342
 Db 174805 GTTGTCTCTCCAGAAATGAAGCTGTAAAGCTTAAAGTGTCTTCAATAGTGTCTTCAATGGAACCCCA 174864
 QY 343 GATGAGTCCATGACTTAAATCTACGTGGACCCCTGGAGCCGCTGTAGACTATGCTC 402
 Db 174865 GATGAGTCCATGACTTAAATCTACGTGGACCCCTGGAGCCGCTGTAGACTATGCTC 174924
 QY 403 TGATGTTAATGACTTGAATGATCACCCTCCGAGGAAATCTCAACTGCAACACCCCTACT 462
 Db 174925 CAATGTTAATGACTTGAATGATCACCCTCCGAGGAAATCTCAACTGCAACACCCCTACT 174984
 QY 463 ACACTCCAAATTCAGTAGGAGGAGTGTAGAGCAGTGTGTAGCCCAACCTCCCAACAGTACT 522
 Db 174985 ATGCCCAATTCAGTAGGAGGAGTGTAGAGCAGTGTGTAGCCCAACCTCCCAACAGTACT 175044
 QY 523 TGGGTTTTCTGTTGAGAGGTTGAGTGTAGAGCAGGACTAGCTGGATTTCTTAGGCTGA 582
 Db 175045 TGGGTTTTCTGTTGAGAGGTTGAGTGTAGAGCAGGACTAGCTGGATTTCTTAGGCTGA 175104
 QY 583 CTAAGAAATCCNAAAGCTTANCTGGGAAGGTGACCGATCCATCTTTTAAACATCGGGCTTG 642
 Db 175105 CTAAGAAATCCNAAAGCTTANCTGGGAAGGTGACCGATCCATCTTTTAAACATCGGGCTTG 175164
 QY 643 CAATTTAGCTCACACCCGACCAATCAGAGAGTCTACTAAATAGCTTAATCAGGCAAAACA 702
 Db 175165 CAATTTAGCTCACAC---CCAAACAGAGAGTCTACTAAATAGCTTAATCAGGCAAAACA 175220
 QY 703 GGAGGTAAGCAATAGCCCAATCATCTATTGCTGTAGAGCAGCAGCGGAGGAGCAAGGATT 762
 Db 175221 GGAGGTAAGCAATAGCCCAATCATCTATTGCTGTAGAGCAGCAGCGGAGGAGCAAGGATC 175280
 QY 763 GGGATATAACTCAGGCAATTCAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCATT 822
 Db 175281 AGGATATAACCCAGGCAATTCAGGCTGGCAACCCCTTTGGGTCCTCCCTCCATT 175340
 QY 823 GTATGGAGCTCTGTTTCACTCTATTCACTCTATTAAATCATCAACTGCA 875

Db 175341 GTATGGAGCTCTGTTTCA--CTGTTTCACTCTATTAATCTTGCAACTGCA 175391
|||||
RESULT 34
US-10-741-600-17995
; Sequence 17995, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17995
; LENGTH: 283351
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)_(283351)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17995
Query Match 36.0%; Score 478.6; DB 19; Length 283351;
Best Local Similarity 90.1%; Pred. No. 1.3e-136;
Matches 534; Conservative 1; Mismatches 52; Indels 6; Gaps 2;
QY 283 GTTGTCTCTCCAGAAATGAAGCTGTAAAGCTAAGAGTACAAATAGTTCTTCAAAATGGACCCCA 342
Db 174805 GTTGTCTCTCTCCAGAAATGAAGCTGTAAAGCTAAGAGTACAAATAGTTCTTCAAAATGGACCCCA 174864
QY 343 GATGAGTCCATGACTGTAAATCTACCTGTGACCCCTGGACCGGCTGTAGACTATGCTC 402
Db 174865 GATGAGTCCATGACTGTAAATCTACCTGTGACCCCTGGACCGGCTGTAGACTATGCTC 174924
QY 403 TGATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAATCTCACTGACACACCCCTACT 462
Db 174925 CAATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAATCTCACTGACACACCCCTACT 174984
QY 463 ACATCTCAATTCAGTAGGAGCAGTGTAGAGCAGTGTGACGCAACCTCCCAACAGTACT 522
Db 174985 ATGCCCCAATTCAGTAGGAGCAGTGTAGAGCAGTGTGACGCAACCTCCCAACAGTACT 175044
QY 523 TGGGTTTTCTCTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGA 582
Db 175045 TGGGTTTTCTCTTGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGA 175104
QY 583 CTAAGAATCCNNAAGCTTANCTGGGAAGGTGACCGATCCATCTTTTAAACATGGGGCTTG 642
Db 175105 CTAAGAATCCCTAAGCTTAGCTGGGAAGGTGACCGATCCATCTTTTAAACATGGGGCTTG 175164
QY 643 CAATCTAGCTACACCCGACCAATCAGAGAGCTCACTAAATGCTAAATCAGGCAAAAACA 702
Db 175165 CAATCTAGCTACAC----CCAAACAGAGAGTTCATAAAATGCTAAATAGGCAAAAATA 175220
QY 703 GGAGGTAAAGCAATAGCCCAATCATCTATTGCTGTAGAGCAGACGCGGAGGACAGAGATT 762
Db 175221 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGTAGAGCAGACGCGGAGGACAGAGATC 175280
QY 763 GGGATATAAATCTAGGCAATTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTT 822
Db 175281 AGGATATAAACCAGGCAATTCGAGCTGGCAACGCAACCCCTTTGGGTCCCTCCCTTT 175340
QY 823 GTATGGAGCTCTGTTTCACTCTATTTCACTCTATTTAAATCATGCAACTGCA 875
Db 175341 GTATGGAGCTCTGTTTCA--CTGTTTCACTCTATTAATCTTGCAACTGCA 175391
|||||
RESULT 35

US-10-719-993-6815
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719.993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)_(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815
Query Match 36.0%; Score 478.6; DB 18; Length 1980090;
Best Local Similarity 90.1%; Pred. No. 3.8e-136;
Matches 534; Conservative 1; Mismatches 52; Indels 6; Gaps 2;
QY 283 GTTGTCTCTCTCCAGAAATGAAGCTGTAAAGCTAAGAGTACAAATAGTTCTTCAAAATGGACCCCA 342
Db 1777083 GTTGTCTCTCTCCAGAAATGAAGCTGTAAAGCTAAGAGTACAAATAGTTCTTCAAAATGGACCCCA 1777142
QY 343 GATGAGTCCATGACTGTAAATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTC 402
Db 1777143 GATGAGTCCATGACTGTAAATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTC 1777202
QY 403 TGATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAATCTCACTGACACACCCCTACT 462
Db 1777203 CAATGTTAATGACATTTGAAGTCAACCCCTCCGAGGAATCTCACTGACACACCCCTACT 1777262
QY 463 ACATCTCAATTCAGTAGGAGCAGTGTAGAGCAGTGTGACGCAACCTCCCAACAGTACT 522
Db 1777263 ATGCCCCAATTCAGTAGGAGCAGTGTAGAGCAGTGTGACGCAACCTCCCAACAGTACT 1777322
QY 523 TGGGTTTTCTCTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGA 582
Db 1777323 TGGGTTTTCTCTTGAGAGGGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGA 1777382
QY 583 CTAAGAATCCNNAAGCTTANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTG 642
Db 1777383 CTAAGAATCCCTAAGCTTAGCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTG 1777442
QY 643 CAATCTAGCTACACCCGACCAATCAGAGAGCTCACTAAATGCTAAATCAGGCAAAAACA 702
Db 1777443 CAATCTAGCTACAC----CCAAACAGAGATTCATTAATAGGCAAAAATA 1777498
QY 703 GGAGGTAAAGCAATAGCCCAATCATCTATTGCTGTAGAGCAGACGCGGAGGACAGAGATT 762
Db 1777499 GGAGGTACAGAAATAGCCCAATCATCTATTGCTGTAGAGCAGACGCGGAGGACAGAGATC 1777558
QY 763 GGGATATAAATCTAGGCAATTCAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTT 822
Db 1777559 AGGATATAAACCAGGCAATTCGAGCTGGCAACGCAACCCCTTTGGGTCCCTCCCTTT 1777618
QY 823 GTATGGAGCTCTGTTTCACTCTATTTCACTCTATTTAAATCATGCAACTGCA 875
Db 1777619 GTATGGAGCTCTGTTTCA--CTGTTTCACTCTATTAATCTTGCAACTGCA 1777669
|||||
RESULT 36
US-10-741-600-17676
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001499
;; CURRENT APPLICATION NUMBER: US/10/741,600
;; CURRENT FILING DATE: 2003-12-22
;; NUMBER OF SEQ ID NOS: 73997
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 17676
;; LENGTH: 1980090
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; NAME/KEY: misc feature
;; LOCATION: (1)...(1980090)
;; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

Query Match 36.0%; Score 478.6; DB 19; Length 1980090;
Best Local Similarity 90.1%; Pred. No. 3.8e-136;
Matches 534; Conservative 1; Mismatches 52; Indels 6; Gaps 2;
QY 283 GTTTGTCTTTCAGAAATTGAAGCTGTAAAGCTACAAATAGTCTTTCAAATGGAAACCCCA 342
Db 1777083 GTTTGTCTTTCAGAAATTGAAGCTGTAAAGCTACAAATAGTCTTTCAAATGGAAACCCCA 1777142
QY 343 GATGAGTCCATGACATAAATCTACCGTGGACCCCTGGACCGCTGCTAGACTATGCTC 402
Db 1777143 GATGAGTCCATGACTAAGATCCACCGCAGACCCCTGGACCGCTGCTAGACTATGCTC 1777202
QY 403 TGATGTTAATGACATTTGAAGTCAACCCCTCCCGAGGAAATCTCAACTGCAACACCCCTACT 462
Db 1777203 CAATGTTAATGACATGAGAGGACCCCTTCGGAGGAAATCTCAACTGCAACACCCCTACT 1777262
QY 463 ACACCTCAATTCAGTAGGAGAGTGTAGAGAGTGTCTAGCCAACTCCGCCAACACACTACT 522
Db 1777263 ATGCCCAATTCAGCAGGAAGCTGTTAGAGCAGTCAATGGCCAACTCCGCCAACACACT 1777322
QY 523 TGGGTTTCTGTTGAGAGGTGGACTGAGAGACAGACTAGCTGCAATTTCTTAGGCTGA 582
Db 1777323 TGGGTTTCTGTTGAGAGAGGGGACTGAGAGACAGACTAGCTGCAATTTCTTAGGCTGA 1777382
QY 583 CTAAGAAATCCNAAGCTTANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTG 642
Db 1777383 CTAAGAAATCCCTAAGCTTAGCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTG 1777442
QY 643 CAATTTAGTTCACACCCGACCAATCAGAGAGTCTCTAAATGCTAATCAGGCAAAAACA 702
Db 1777443 CAATTTAGTTCACAC---CCAAACAGAGAGTTCCTAAATGCTAATTTAGGCAAAAATA 1777498
QY 703 GAGGTTAAGCAATAGCCATCATCTATTGCTGAGAGACAGCGGAGGAGCAAGGATT 762
Db 1777499 GAGGTTAAGCAATAGCCATCATCTATTGCTGAGAGACAGCGGAGGAGCAAGGATC 1777558
QY 763 GGGATATAAATCAGGCAATCAAGCCAGACAAAGCAACCCCTTTGGGTCCCTCCCAATT 822
Db 1777559 AGGATATAAACCCAGGCAATTCAGAGTTCGGCAACCGCAACCCCTTTGGGTCCCTCCCTTT 1777618
QY 823 GTATGGAGCTCTGTTTTCATCTATTTCATCTATTAAATCATGCAACTGCA 875
Db 1777619 GTATGGAGCTCTGTTTTCATCTATTAAATCTTTGCAACTGCA 1777669

RESULT 37
US-10-276-774-678
; Sequence 678, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18

;; PRIOR APPLICATION NUMBER: 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 2700
;; SOFTWARE: Custom
;; SEQ ID NO 678
;; LENGTH: 2052
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-276-774-678

Query Match 35.3%; Score 468.8; DB 17; Length 2052;
Best Local Similarity 88.5%; Pred. No. 1.1e-134;
Matches 533; Conservative 0; Mismatches 59; Indels 10; Gaps 2;
QY 257 CCCTGTATCTTCAACTTCTTGTAAAGTTTGTCTTTCAGAAATTCAGAGCTCTTAAGCTA 316
Db 993 CCCTGTATCTTAAACCTCTTGTAAAGTTTGTCTTTCAGAAATTCAGAGCTCTTAAGCTA 1052
QY 317 CAAATAGTCTTCAAAATGGAACCCAGATGCAATGCACTTAAATCTTACCGTGGACCC 376
Db 1053 CACATGCTTCTTCAAAATGGAACCCAGATGCAATGCACTTAAATCTTACCGGATCC 1112
QY 377 CTGGACCGCTCTAGACTATGCTCTGATGTTAATGACATTTGAAGTCAACCCCTCCCGAG 436
Db 1113 CTGGACCGCTCTAGCCATGCTCCGATGTTAATGACATGGAAGGCACTCTCTCCCGAG 1172
QY 437 GAAATCTCACTGCAACACCCCTACTACACTCCAATTCAGTAGGAGAGCTTAGAGCAGT 496
Db 1173 GAAATCTCAAAATGCAACACCCCTACTATGCCCAATTAAGCAGGAGAGCTTAGAGCAGT 1232
QY 497 TGTGAGCCAACTCCCCCAACAGTACTTGGGTTTTCTGTTGAGAGGCTGCACTGAGAGAC 556
Db 1233 CAGAGCCCAACCTCCCCCAACAGCCTTGGGTTTTCTGTTGAGAGGCTGCACTGAGAGAC 1292
QY 557 AGGACTAGCTGGATTTCCTTAGCTGACTAAGAAATCCCAAGCTTANCTGGGAAGGTGACC 616
Db 1293 AGGACTAGCTGGATTTCCTTAGCCGGATAAGAAATCCCTAAGCTTAGCTGGGAAGGTGACT 1352
QY 617 GCATCATCTTTAAACATGGGGCTTGCACACTTAGCTCACCACCGCAATC----- 667
Db 1353 GTATCCACCTTTAAACACAGGGCTTGCACACTTAGCTCACCACCGCAATCAGGTAGTAA 1412
QY 668 AGAGAGCTCACTTAAATGCTTAATCAGGCAAAAACAGAGGTAAGCAATAGCCAATCATC 727
Db 1413 AGAGAGCTCACTTAAATGCTTAATTAGGCAAAAACAGAGGTAAGCAATAGCTAAT-ATC 1471
QY 728 TATTGCTGAGAGCAGCGGGAAGGACAAAGGATTGGGATATAAACTCAGGCAATTCAGAGC 787
Db 1472 TATCAGCTGAGAGTACAGGGGAGGAGCAATGATTGGGATAGAAACCCAGGCAATTCGAGC 1531
QY 788 CAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTTATGGGAGCTCTGTTTTCACCTCA 847
Db 1532 CGGCAACGGCAACCCCTTTGGGTCTCTCTTCCATTTTATGGGAGCTCTGTTTTCACCTCA 1591
QY 848 TT 849
Db 1592 TT 1593

RESULT 38
US-10-632-793-28
; Sequence 28, Application US/10632793
; Publication No. US20040048298A1
; GENERAL INFORMATION:
; APPLICANT: PARANHOS-BACCALA, Glaucia
; APPLICANT: MALLET, Francois
; APPLICANT: VOISSET, Cecile
; TITLE OF INVENTION: ENDOGENOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
; TITLE OF INVENTION: AUTOIMMUNE DISEASE, LABELING METHOD AND REAGENT
; FILE REFERENCE: 110048
; CURRENT APPLICATION NUMBER: US/10/632,793

Fri Feb 25 16:26:28 2005

```
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/869,927
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/FR00/00144
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: FR 99/00888
; PRIOR FILING DATE: 1999-01-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-632-793-28

Query Match      35.0%; Score 465; DB 17; Length 3372;
Best Local Similarity 89.7%; Pred. No. 2.2e-133;
Matches 498; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 321 TAGTCTTCAATGGAACCCAGATGCGATCCATGACTAAATCTACGGTGAGCCCTGG 380
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2797 TAAAACTACAATGGAGCCCAAGATGCGATCCAAAGACTTAAGATCTACCGCAGACCCCTGG 2856
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 381 ACCGCCCTGTAGACTATGCTCTGATGTTATGATGACATTCACCCCTCCCGAGGAAA 440
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2857 ACCGCCCTGTAGCCCAAGATCTGATGTTATGATGACATTCAAAGGCACCCCTCTCGAGGAAA 2916
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 441 TCTCAACTGCGCAACCCCTACTACACTCCAATTCAGTAGGAGCAGTTCAGAGCAGTTGTC 500
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2917 TCTCAGCTGCGCAACCTCTACTAGCCGCCAATTCAGCAGGAAGCAGTTAGAGCGGTGTC 2976
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 501 AGCCAACTCCCAACAGTACTTGGGTTTCTGTTGAGAGGGTGGACTGAGAGCAGGA 560
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2977 GCGCAACCTCCCAACAGCAGTCTAGGTTTCTGTTGAGTGGGGAGCTGAGAGCAGGA 3036
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 561 CTAGCTGGAATTCCTAGGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGACCGCAT 620
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3037 CTAGCTGGAATTCCTAGGCTGATTAAGAAATCCCTAAGCTAGCTGGGAAGGTGACCAT 3096
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 621 CCATCTTTAAACATGGGCTTGCAACTTAGCTCACAACCGCAACCAATCAGAGGCTCACTA 680
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3097 CCACCTTTAAACACAGGGCTTGCAACTTAGCTCACAACCTGACCAATCAGAGGCTCACTA 3156
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 681 AAATGCTAATCGGCAAAACAGGAGCTTAAGCAATAGCAATCATCTATTGCTGAGAG 740
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3157 AAATGCTAATAGGCAAAACAGGAGGTAAGAAATAGCAATCATCTATTGCTGAGAG 3216
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 741 CACAGCGGAAGGACAAGATTTGGGATATAAATCAGGCATTCAGCCAGCAACAGCAAC 800
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3217 CACAGCAGAGGGAACATGATCGGGATATAACCCCAAGTTTTCGAGCCGCAACCGCAAC 3276
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 801 CCCCTTTGGGTCCTCCCATCTGATGGAGGCTCTGTTTCACTCTATTTCATCTATTA 860
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3277 CCCCTTTGGGTCCTCCCTCTGTTATGGAGGCTCTGTTTCACTATTTCATCTATTA 3336
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 861 AATCATGCAACTGCA 875
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 3337 AATCTGCAACTGCA 3351
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 39
US-09-997-722-148
; Sequence 148, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 22436
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-997-722-148

Query Match      35.0%; Score 465; DB 11; Length 22436;
Best Local Similarity 88.8%; Pred. No. 6e-133;
Matches 539; Conservative 0; Mismatches 57; Indels 11; Gaps 3;

Qy 257 CCCTGTATCTTCAACTCTCTTGAAGTTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 316
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5655 CCCTGTACCTTTAACTCTCTTGAAGTTTGTCTCTCCAGAAATCGAAGCTGTAAAGCTA 5714
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 317 CAAATAGTTCTTCAAAATGGAACCCAGATGCGATCCATGACTAAAATCTTACCGTGGACCC 376
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5715 CAAATGGTTCTTCAAAATGGAACCCAGATGCGATCCATGACTAAAATCTTACCGTGGACCC 5774
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 377 CTGGAACGGGCTGCTAGACT-ATGCTCTGATGTTAATGACATTTGAAGTCACTCCCGA 435
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5775 CTGGAACGGGCTGCTAGCTCTATGCTCCAAATGTTAATGACATTTGAAGGCAACTCTCCCGA 5834
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 436 GGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGGAGCAGTTAGAGCAG 495
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5835 GGAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGGAGCAGTTAGAGCAG 5894
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 496 TTGTACGCCCAACCTCCCCCAACAGTACTTTGGGTTTTCTGTTGAGAGGTTGACTGAGAGA 555
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5895 TCATTGGCCCAACCTCCCCCAACAGTACTTTGGGTTTTCTGTTGAGAGGTTGACTGAGAGA- 5953
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 556 CAGGACTAGCTGGAATTTCTAGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGAC 615
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5954 CAGGACTAGCTGGAATTTCTAGCTGACTAAGAAATCCNAAAGCTTANCTGGGAAGGTGAC 6013
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 616 CGCATCTCATCTTTAAACATGGGCTTGCAACTTAGCTCACCACCGGCAATC----- 667
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6014 TGACCCACCTTTAAACACGGGGCTTGCAACTTAGTTTACACCTGACCAATCAAGTAGTA 6073
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 668 -AGAGAGCTCACTAAAATGCTAATCAGGCAAAAACAGAGGTTAAGCAATAGCCAAATCAT 726
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6074 AAGAGAGCTCACTAAAATGCTAATTAGGCAAAAACAGAGGTTAAGCAATAGCCAAATCAT 6133
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 727 CTATTGCTGAGACACAGCGGAGGCAAGGATTTGGGATATAAACTCAGGCAATTCAGG 786
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6134 CTATCCTGAGAGCAAGCGTGGGGGCAATGATCAGGATATAAACCCAGGCAATTTGAG 6193
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 787 CCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCAATTTGATGGGAGCTCTGTTTTCATCT 846
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6194 CCAGCAATGCTACCTCTTTGGGTCCTCCCTCCCTCCCTTTGATGGGAGCTCTGTTTTCATCT 846
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 847 ATTTTCA 853
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 6254 ATTTTCA 6260
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 40
US-10-087-192-1666
; Sequence 1666, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1666
; LENGTH: 285020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(285020)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1666

Query Match      34.6%; Score 460; DB 13; Length 285020;
Best Local Similarity 88.6%; Pred. No. 8.5e-131;
Matches 535; Conservative 0; Mismatches 57; Indels 12; Gaps 3;

QY 257 CCCTGTATCTTCAACTTCCTTGTAAAGTTTGTCTTCCAGAAATTGAAGCTGTAAAGCTA 316
Db 279486 CCCTGTATCTTCAACTTCCTTGTAAAGTTTGTCTTCCAGAAATTGAAGCTGTAAAGCTA 279545

QY 317 CAAATAGTTCCTTCAATGAACCCAGATGCAGTCCATGACTTAAATCTACCGTGACCC 376
Db 279546 CAAATAGTTCCTTCAATGAACCCAGATGCAGTCCATGACTTAAATCTACCGTGACCC 279605

QY 377 CTGGACCGGCTGTAGACTATGCTCTGATGTATGATTAATGACATTGAAGTCAACCCCTCCCGAG 436
Db 279606 CTGGACCGGCTGTAGACTATGCTCTGATGTATGATTAATGACATTGAAGTCAACCCCTCCCGAG 279665

QY 437 GAAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGGAGCAGTTAGAGCAGT 496
Db 279666 GAACTCTCAACTGCACGACCCCTACTACTATGCCCCAATTCAGCAGGAGCAGTTAGAGTGT 279725

QY 497 TGTACGCCAACCTCCCAACAGTACTTGGGTTTTCCTGTTGAGAGGGTGGACTGAGAGAC 556
Db 279726 CATTGSCCAACCTCCCAACAGCAGCTTGGGTTTTCCTGTTGAGAGTGGGAGCTGAGAGAC 279785

QY 557 AGGACTAGCTGGATTTCCCTAGGCTGACTAAG--AATCCCNAGCCCTANCTGGGAGGTGA 614
Db 279786 AGGACTAGCTGGATTTCCCTAGGCTGACTAAGAAATCCCTAAGCCCTAGCTGGGAGGTGA 279845

QY 615 CCGCATCCATCTTTAAACATGGGCTTGCAACTTAGCTCACACCCGACCAATC----- 667
Db 279846 CCGCATCCACCTTTAAACAGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGGGAGT 279905

QY 668 --AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGAGGTTAAAGCAATAGCCCAATCA 725
Db 279906 AAAGACAGCTCACTAAATGCTAATTAGGCTAAACAGAGGTTAAAGAAATAGCCCAATCA 279965

QY 726 TCTATTGCTGAGACACAGCGGAGGACAGGATTTGGATATAACTCAGGCATTCAA 785
Db 279966 TCTATCGCTGAGAGCAC--GAGGGAGGGAACAATGATCAGGATATAAAACCCAGGCATTCAA 280024

QY 786 GCCAGCAACAGAACCCCTTTGGGTCCCTCCATTGATGGAGCTCTGTTTTCACTC 845
Db 280025 GCCGCAGTGGCTACCCGCTTTGGGTCCCTCTGTTTGTATGGAGCTCTGTTTCACTC 280084

QY 846 TATT 849
Db 280085 TATT 280088
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2005, 06:52:54 ; Search time 4132.82 Seconds
(without alignments)
12240.400 Million cell updates/sec

Title: US-09-319-156B-12
Perfect score: 1329
Sequence: 1 tcaaatgaagagcttttag.....gcaaggatccccagtaaca 1329

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gsea.*
- 9: gb_gsea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	706.8	53.2	2748	CR605851	full-leng
2	702.8	52.9	2749	CR617248	full-leng
3	702	52.8	2500	CR622175	full-leng
4	699	52.6	2748	CR625046	full-leng
5	684.2	51.5	1071	CR635066	full-leng
6	678.6	51.1	2716	CR613169	full-leng
7	669.4	50.4	998	CR633769	full-leng
8	661.2	49.8	1058	CR637830	full-leng
9	652.2	49.1	1019	CR639636	full-leng
10	610	45.9	966	CR638017	full-leng
11	608.2	45.8	758	CR635720	full-leng
12	593.4	44.7	771	CR618786	full-leng
13	592	44.5	4204	CR635153	full-leng
14	561	42.2	719	CR627239	full-leng
15	544	40.9	900	CR645915	full-leng
16	500.6	37.7	1002	CR635382	full-leng
17	480.8	36.2	679	CR611369	full-leng
18	480.8	36.2	999	CR618330	full-leng
19	475.2	35.8	683	CR613452	full-leng
20	474	35.7	669	CR604953	full-leng
21	473	35.6	681	CR611697	full-leng
22	469.6	35.3	710	CR609634	full-leng
23	468.2	35.2	1160	CR630968	full-leng
24	465.2	35.0	893	CR643792	full-leng

25	456.2	34.3	670	9	AG072852	Pan trogl
26	455	34.2	712	9	AG099717	Pan trogl
27	439.2	33.0	689	9	AG121669	Pan trogl
28	438.8	33.0	921	8	BH149565	ENTOG48TR
29	435.6	32.8	679	9	AG076758	Pan trogl
30	433.6	32.6	683	9	AG102951	Pan trogl
31	433.6	32.4	611	8	AG081711	RPC111-16
32	429.4	32.3	736	9	AG086046	Pan trogl
33	427.6	32.2	714	8	AG023872	RPC111-64
34	424.4	31.9	697	9	AG108186	Pan trogl
35	423.8	31.9	675	9	AG117095	Pan trogl
36	422.2	31.8	797	6	CB990962	AGENCOURT
37	422.2	31.8	803	6	CB992422	AGENCOURT
38	422.2	31.8	826	6	CB990802	AGENCOURT
39	422.2	31.8	837	6	CB996320	AGENCOURT
40	422.2	31.8	852	6	CB994833	AGENCOURT
41	419.4	31.6	672	9	AG040241	Pan trogl
42	418.8	31.5	611	7	CR542951	DKF29470A
43	418.2	31.5	870	7	CN645411	ILLUMINIGEN
44	416.2	31.3	692	9	AG114407	Pan trogl
45	415.6	31.3	494	1	AA781423	aj26c03.s

ALIGNMENTS

RESULT 1
CR605851
LOCUS full-length cDNA clone CS0DE012YJ24 of Placenta of Homo sapiens (human).
DEFINITION
ACCESSION CR605851
VERSION CR605851.1 GI:50486658
KEYWORDS HTC; CINSIT_CDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2748)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 2748)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
source
1..2748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YJ24"
/tissue_type="placenta"
/plasmid="pCMVSPORT_6"

Query Match 53.2%; Score 706.8; DB 3; Length 2748;
Best Local Similarity 90.5%; Pred. No. 1.2e+202;
Matches 769; Conservative 0; Mismatches 69; Indels 12; Gaps 1;

QY 1 TCAAAATCGAAGAGCTTTAGACTTGTACCGCCAAAGAGGGGACCTGTTATTTT 60
|||||
Db 1911 TCAAAATCGAAGAGCTTTAGACTTGTACCGCTGAAGAGGGGACCTGTTATTTT 1970
|||||

61 AGGGGAAGAAGTCTGTAGTATGTTAAATCAATCTGGAATCAATTAAGTGAAGAAGTTAAAGA 120
 1971 AGGGGAAGAAGTCTGTATTAATGTTAAATCAATCGGAATCGTCACTGAGAAGAAGTTAAAGA 2030
 121 AATTGAGATCGAATTAATAGTAGAGCAGAGGACCTTCAAAACACTGCACACCTGGGCGCT 180
 2031 AATTGAGATCGAATTAACACGCTAGAGCAGAGGAGCTTCGAAACACTGCACACCTGGGCGCT 2090
 181 CCTCAGCCAAATGAGTGGCTGAGCTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 240
 2091 CTTAGCCAAATGAGTGGCTGAGTCTCCCTTCTTAGGACCTCTAGCAGCTATAATATT 2150
 241 TTTACTCTCTTTGGACCGCTGATCTTCAACTCTCTGTTAAGTTTGTCTCTTCAGAAAT 300
 2151 GCTACTCTCTTTGGACCGCTGATCTTTAACTCTCTGTTAAGTTTGTCTCTTCAGAAAT 2210
 301 TGAAGCTGTAAGCTTACAATAGTCTTCAATGGAACCCAGATGCACTGCATGACTAA 360
 2211 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCACTGCATGACTAA 2258
 361 AATCTACCGTGGACCGCTGGACCGGCTGTAGACTATGCTCTGATGTTAATGACTTGA 420
 2259 GATCTACCGCAGACCGCTGGACCGGCTGTAGCCACAGATCTGATGTTAATGACTTGA 2318
 421 AGTACCGCTCCCGAGGAAATCTCACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
 2319 AGGCACCGCTCCCGAGGAAATCTCAGCTGCAACCGCTCTACTACCGCCCAATTCAGTAGG 2378
 481 AAGCAGTTAGACAGTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 540
 2379 AAGCAGTTAGACAGTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTTGAGA 2438
 541 GGGTGGACTCAGAGCAGACTAGCTGGATTTCTAGGCTGCTAAGAAATCCNAAAGCT 600
 2439 TGGGGACTCAGAGCAGACTAGCTGGATTTCTAGGCTGCTAAGAAATCCNAAAGCT 2498
 601 ANCTGGGAAGGTGACCGCATTCATTTTAAACATGGGGCTTGCACCTAGCTCAGACCG 660
 2499 AGCTGGGAAGGTGACCACTCAACCTTTAAACACGGGGCTTGCACCTAGCTCAGACCTG 2558
 661 ACCAATCAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
 2559 ACCAATCAGAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCC 2618
 721 AATCATCTATTGCTCAGAGCAGCGGAAAGCAAGGATTTGGGATATAAATCAGGCA 780
 2619 AATCATCTATTGCTCAGAGCAGCGGAGGCAATGATCGGGATATAAATCAGGCA 2678
 781 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGATGGGAGCTCTGTTT 840
 2679 TTCAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCTTTGATGGGAGCTCTGTTT 2738
 841 CACTCTATT 850
 2739 CATGCTATT 2748

RESULT 2
 LOCUS CR617248
 DEFINITION full-length cDNA clone CS0D1022YJ18 of Placenta Cot 25-normalized of Homo sapiens (human).
 ACCESSION CR617248
 VERSION CR617248.1 GI:50498055
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2749)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalized

JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 2749)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 COMMENT - Web : www.genoscope.cns.fr)
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 source
 1. 2749
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1022YJ18"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

Query Match 52.9%; Score 702.8; DB 3; Length 2749;
 Best Local Similarity 90.4%; Pred. No. 1.9e-201;
 Matches 765; Conservative 0; Mismatches 69; Indels 12; Gaps 1;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTTAAACCGCTGAAAGAGGGGAACTGTTTATTTT 60
 Db 1916 TCAAAATCGAAGAGCTTTAGACTTGTCTTAAACCGCTGAAAGAGGGGAACTGTTTATTTT 1975
 QY 61 AGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCATTAATCAGAAAGTTAAAGA 120
 Db 1976 AGGGAAGATGCTGTTAGTATGTTAATCAATCTGGAATCATTAATCAGAAAGTTAAAGA 2035
 QY 121 AATTGAGATCGAATTAATAGTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGCGCT 180
 Db 2036 AATTGAGATCGAATTAATAGTAGAGCAGAGGACCTTCAAAACACTGCACCTGGGCGCT 2095
 QY 181 CCTCAGCAATGAGATGCTTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 240
 Db 2096 CCTCAGCAATGAGATGCTTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATAATTT 2155
 QY 241 TTTACTCTCTTTGGACCGCTGATCTTTAACTCTCTGTTAAGTTTGTCTCTCCAGAAAT 300
 Db 2156 GCTACTCTCTTTGGACCGCTGATCTTTAACTCTCTGTTAAGTTTGTCTCTCCAGAAAT 2215
 QY 301 TGAAGCTGTAAGCTCAAAATAGTTCTTCAATGGAACCCAGATGCACTGCATGACTAA 360
 Db 2216 CGAAGCTGTAAGCTCAAAATAGTTCTTCAATGGAACCCAGATGCACTGCATGACTAA 2263
 QY 361 AATCTACCGTGGACCGCTTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTTGA 420
 Db 2264 GATCTACCGCAGACCGCTTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACTTGA 2323
 QY 421 AGTACCGCTCCCGAGGAAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
 Db 2324 AGGACCGCTCTCTGAGGAAATCTCAGCTGCAACCTCTACTACGCCCAATTCAGTAGG 2383
 QY 481 AAGCAGTTAGAGCAGTGTGTAGCCAACTCTCCCAACAGTACTTGGGTTTCTGTTGAGA 540
 Db 2384 AAGCAGTTAGAGCAGTGTGTAGCCAACTCTCCCAACAGTACTTGGGTTTCTGTTGAGA 2443
 QY 541 GGGTGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCT 600
 Db 2444 TGGGGGACTGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCNAAAGCT 2503
 QY 601 ANCTGGGAAGGTGACCGCATCCATCTTTTAAACATGGGGCTTGCACCTTAGCTCACACCG 660
 Db 2504 AGCTGGGAAGGTGACCAATCCACCTTTTAAACACGGGGCTTGCACCTTAGCTCACACCTG 2563

QY 661 ACCAATCAGAGGCTCAATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
 Db 2564 ACCAATCAGAGGCTCAATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2623

QY 721 AATCATCTATTGCTTGGCTGAGAGCACAGCGGAGGACAGGATTTGGGATATAAATCAGGCA 780
 Db 2624 AATCATCTATTGCTTGGCTGAGAGCACAGCGGAGGACAGGATTTGGGATATAAATCAGGCA 2683

QY 781 TTCAAGCCAGCAGACAGCAACCCCTTTGGGTGCTCCCTCCATGTTATGGAGCTCTGTTTT 840
 Db 2684 TTCAAGCCAGCAGACAGCAACCCCTTTGGGTGCTCCCTCCATGTTATGGAGCTCTGTTTT 2743

QY 841 CACTCT 846
 Db 2744 CATGCT 2749

RESULT 3
 CR622175
 LOCUS full-length cDNA clone CS0D1051YM13 of Placenta Cot 25-normalized
 DEFINITION of Homo sapiens (human).
 ACCESSION CR622175
 VERSION CR622175.1 GI:50502982
 KEYWORDS HTC; CNSLT_CDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2500)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 2500)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES Location/Qualifiers
 source 1..2500
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1051YM13"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN
 Query Match 52.8%; Score 702; DB 3; Length 2500;
 Best Local Similarity 90.6%; Pred. No. 3.2e-201;
 Matches 763; Conservative 0; Mismatches 67; Indels 12; Gaps 1;
 QY 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATCTACTAGAGAAAGTTAAAGA.120
 Db 1669 TCAAAATCGAAGAGCTTTAGACTTGTCTAATCAATCTGGAATCATCTACTAGAGAAAGTTAAAGA 1728

QY 61 AGGGAGAGATGCTGTTAGTATGTTAATCAATCTGGAATCATCTACTAGAGAAAGTTAAAGA.120
 Db 1729 AGGGAGAGATGCTGTTAGTATGTTAATCAATCTGGAATCATCTACTAGAGAAAGTTAAAGA 1788

QY 121 AATTTCAGATCGAATATAATGTTAGAGCAGAGGACCTTTCAAACAACCTGCACCTCGGGCCCT 180
 Db 1789 AATTTCAGATCGAATATAACAGTAGAGAGGAGCTTCGAAACACACTGGACCTCGGGCCCT 1848

QY 181 CCTCAGCAATGGATGCGCTGGACCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
 Db 1849 CCTCAGCAATGGATGCGCTGGATTTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 1908

QY 241 TTTTACTCTCTTTGGAGCCCTGTATCTTCAACTCTCTTGTAAAGTTTGTCTCTCCAGAAAT 300
 Db 1909 GCTACTCTCTCTTTGGAGCCCTGTATCTTAACTCTCTTGTAACTTTGTCTCTCCAGAAAT 1968

QY 301 TGAAGCTGTAAAGCTCAAAATAGTTCTTCAATAGAAACCCAGATGCGAGTCAATGACTAA 360
 Db 1969 CGAAGCTGTAAACTA-----CAAAATGGAGCCCAAGATGCGAGTCAAGACTAA 2016

QY 361 AATCTACCGTGGACCCCTGGACCGGCTGTAGACTATGCTGTGATGTTAATGACATTGA 420
 Db 2017 GATCTACCGCAGACCCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCA 2076

QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGACACACCCCTACTACACTCCAATTCAGTAGG 480
 Db 2077 AGGCACCCCTCTGAGGAAATCTAGCTGACACACCTCTACTACGCCCAATTCAGCAGG 2136

QY 481 AAGCAGTTAGAGCAGTTGCTCAGCCCAACCTCCCAACAGTACTTGGGTTTTCTCTGTGAGA 540
 Db 2137 AAGCAGTTAGAGCGTCTGCGCCCAACCTCCCAACAGTACTTGGGTTTTCTCTGTGAGA 2196

QY 541 GGGTGGACTGAGAGCAGGACTAGCTGATTTCTTAGCTGACTAGTAAAGTCCCAAGCCT 600
 Db 2197 TGGGGGACTGAGAGCAGGACTAGCTGATTTCTTAGCTGACTAGTAAAGTCCCAAGCCT 2256

QY 601 ANCTGGGAGGTGACCGCATCATCTTTAAACATGGGCTTGCACCTTAGCTCACAACCCG 660
 Db 2257 AGCTGGGAGGTGACCATCCCTTTAAACATGGGCTTGCACCTTAGCTCACAACCTG 2316

QY 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 720
 Db 2317 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCC 2376

QY 721 AATCATCTATTGCTTGGAGACACGCGGAGGACAGGATTTGGATATAAATCAAGTC 780
 Db 2377 AATCATCTATTGCTTGGAGACACGCGGAGGACAGGATTTGGATATAAATCAAGTC 2436

QY 781 TTCAGCCAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTCTTTT 840
 Db 2437 TTCAGCCAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCTCTTTT 2496

QY 841 CA 842
 Db 2497 CA 2498

RESULT 4
 CR625046
 LOCUS full-length cDNA clone CS0D1044YK06 of Placenta Cot 25-normalized
 DEFINITION of Homo sapiens (human).
 ACCESSION CR625046
 VERSION CR625046.1 GI:50505853
 KEYWORDS HTC; CNSLT_CDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2748)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 2748)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

Location/Qualifiers
1..2748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1044YK06"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 52.6%; Score 699; DB 3; Length 2748;
Best Local Similarity 90.6%; Pred. No. 2.7e-200;
Matches 760; Conservative 0; Mismatches 67; Indels 12; Gaps 1;
QY 1 TCATAATCGAGAGCTTTAGACTTGCTAACCGCCAAAGAGGGGGAACCTGTTATTTTT 60
DB 1922 TCATAATCGAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTATTTTT 1981
QY 61 AGGGGAAGATGCTCTAGTAGTGTAAATCAATCTGAATCAATCTAGAGAAAGTTAAAGA 120
DB 1982 AGGGGAAGATGCTCTAGTAGTGTAAATCAATCTGAATCAATCTAGAGAAAGTTAAAGA 2041
QY 121 AATTGAGATCGAATATAATGATGAGCAGAGGAGCACTTCANAAACACTGCACCTGGGGCCT 180
DB 2042 AATTGAGATCGAATATAATGATGAGCAGAGGAGCTTCGAAACACTGCACCTGGGGCCT 2101
QY 181 CCTAGCAATAGGATGCCCTGAGCTCTCCCTCTTAGACCTCTAGACCTCTAGCAGCTATAATT 240
DB 2102 CCTAGCAATAGGATGCCCTGAGCTCTCCCTCTTAGACCTCTAGCAGCTATAATT 2161
QY 241 TTACTCTCTTTGAGACCTGATCTTCAACTTCTCTGTTAAAGTTTGTCTTCCAGAAAT 300
DB 2162 GCTACTCTCTTTGAGACCTGATCTTAACTCTCTGTTAACTTTGTCTTCCAGAAAT 2221
QY 301 TGAAGCTGTAAGCTTACAATAGTTCTTTCAATGGAACCCAGATGCGAGTCCATGACTAA 360
DB 2222 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCGAGTCCAGACTAA 2269
QY 361 AATCTACCGTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420
DB 2270 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 2329
QY 421 AGTCACCCCTCCGAGGAAATCTCAACTGCAACCCCTACTACACTCCAATTCAGTAGG 480
DB 2330 AGGCACCCCTCTGAGGAAATCTAGCTGCAACACCTCTACTACGCCCCCAATTCAGCAGG 2389
QY 481 AAGCAGTTAGACGAGTTGTGACCAACCTCCCAACAGTACTTTGGGTTTTCTGTTGAGA 540
DB 2390 AAGCAGTTAGACGAGTGTGCGCCAACTCCCAACAGCACTTAGTTAGTTTCTGTTGAGA 2449
QY 541 GGGTGGACTGAGACAGCACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
DB 2450 TGGGGGACTGAGACAGCACTAGCTGATTTCTTAGGCTGACTAAGAAATCCCTAAGCCT 2509
QY 601 ANCTGGGAAGGTGACCGCATCATCTTTAAACATGGGGCTTGGCACTTAGCTCACACCCG 660
DB 2510 AGCTGGGAAGGTGACCATCATCACTTTAAACAGGGGCTTGGCACTTAGCTCACACCTG 2569
QY 661 ACCAATCAGAGGCTCACTAAATGCTAATCAGGCAAAAAAGGAGGTAAGCAATAGCC 720
DB 2570 ACCAATCAGAGGCTCACTAAATGCTAATTAGGCAAGACAGGAGGTAAGCAATAGCC 2629
QY 721 AATCATCTATTGCTGAGACAGCAGGAGGACAGGATTTGGGATATAAATCAAGCA 780
DB 2630 AATCATCTATTGCTGAGACAGCAGGAGGACAAATTTGATCGGGATATAAACCAGTC 2689
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCCCTCCCATTTGTTATGGGAGCTCTGTTT 839

DB 2690 TTCAGCGGCAACGGCAACCCCTTTGGGTCCCTTCCCTTTGTATGGGAGCTCTGTTT 2748
RESULT 5
BX365066/c
LOCUS
DEFINITION BX365066 Homo sapiens PLACENTA Cot 25-NORMALIZED Homo sapiens cDNA
clone CS0D1046YA18 3-PRIME, mRNA sequence.
ACCESSION BX365066
VERSION BX365066.2 GI:46304105
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1071)
AUTHORS Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 5, 2003 this sequence version replaced gi:30374869.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4215.x
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS1AI012ZC10NPI&c=4215.r.
FEATURES
Location/Qualifiers
1..1071
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1046YA18"
/tissue_type="PLACENTA Cot 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA Cot 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 51.5%; Score 684.2; DB 5; Length 1071;
Best Local Similarity 89.9%; Pred. No. 6.4e-196;
Matches 755; Conservative 3; Mismatches 69; Indels 13; Gaps 2;
QY 1 TCATAATCGAAGAGCTTTAGACTTGCTAACCGCCAAAGAGGGGGAACCTGTTATTTTT 60
DB 827 TCATAATCGAAGAGCTTTAGACTTGCTAACCGCTGAAAGAGGGGGAACCTGTTATTTTT 768
QY 61 AGGGGAAGATGCTGTTAGTAGTGTAAATCAATCTGGAATCAATCTAGAAAGTTAAAGA 120
DB 767 AGGGGAAGATGCTGTTAGTAGTGTAAATCAATCTGGAATCGTCACTGAAAGTTAAAGA 708
QY 121 AATTGAGATCGAATATAATGATGAGCAGAGGAGCTTCANAAACACTGCACCTGGGGCCT 180
DB 707 AATTGAGATCGAATATAATGATGAGCAGAGGAGCTTCGAAACACTGCACCTGGGGCCT 648
QY 181 CCTAGCAATAGGATGCCCTGAGCTCTCCCTCTTAGGACCTCTAGCAGCTATAATT 240
DB 647 CCTAGCAATAGGATGCCCTGAGCTCTCCCTCTTAGGACCTCTAGCAGCTATAATT 588
QY 241 TTTACTCTCTTTGGACCTGATCTTCAACTCTCTTGTAAAGTTTGTCTTCCAGAAAT 300
DB 587 GCTACTCTCTTTGGACCTGATCTTAACTCTTGTAACTTTGTCTTCCAGAAAT 528
QY 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCAAGTCCATGACTAA 360

Db 527 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCAGTCCAAAGACTAA 480

Qy 361 AATCTACCTGGACCCCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATTGA 420

Db 479 GATCTACCGCAGACCCCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCA 420

Qy 421 AGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGG 480

Db 419 AGGCACCCCTCCCTGAGGAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG 360

Qy 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTTGGGTTTTTCTGTTTGA 540

Db 359 AAGCAGTTAGAGCGGTGCTGGGCCAACCTCCCAACAGCACTTAGGTTTTTCTGTTTGA 300

Qy 541 GGTGAGCTGAGAGACAGACTAGCTGGATTCTCTAGGCTGACTAAGATCCCNAGCCT 600

Db 299 TGGGGGACTGAGAGACAGGACTAGCTGGATTCTCTAGGCTGACTAAGATCCCNAGCCT 240

Qy 601 ANCTGGGAAGTGCACCGCATCTCTTTAAACATGGGCTTGCAACTTAGCTCACACCCG 660

Db 239 AGCTGGGAAGTGACCACTCCACATCCACCTTTAAACAGCGGGCTTGCAACTTAGCTCACACCTG 180

Qy 661 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 720

Db 179 ACCAATCAGAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 120

Qy 721 AATCATCTATTGCTGAGACACAGCGGAGAGCAAGGATTTGGGATATAAATCAGGCA 780

Db 119 AATCATCTATTGCTGAGACACAGCGGAGAGCAAGGATTTGGGATATAAATCAGGCA 60

Qy 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGCTCCCTCCCACTTATGAGGAGCTCTGTTTT 840

Db 59 TTCAGCGCGCAACGGCAACCCCTTTGGTNNCCCTCCC-TTGTATGCGASMTCKGTTTT 1

RESULT 6
CR613169
LOCUS 2716 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DE013Y120 of Placenta of Homo sapiens (human).

ACCESSION CR613169
VERSION CR613169.1 GI:50493976
KEYWORDS HTC; CNSLT cDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2716)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2716)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source
1. 2716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013Y120"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 51.1%; Score 678.6; DB 3; Length 2716;
Best Local Similarity 90.5%; Pred. No. 4.3e-194;
Matches 739; Conservative 0; Mismatches 66; Indels 12; Gaps 1;
Qy 1 TCAAAATCGAAGAGCTTTAGAGCTTCTAAACCCCAAGAGGGGGAACCTGTTTATTTT 60
Db 1912 TCAAAATCGAAGAGCTTTAGAGCTTCTAAACCCCTGAAGAGGGGGAACCTGTTTATTTT 1971

Qy 61 AGGGGAAGAACTCTGTTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
Db 1972 AGGGGAAGAACTCTGTTATTTAATCAATCCGGAATCGTCACTGAGAAAGTTAAAGA 2031

Qy 121 AATTTGAGATCGAATATATATGATAGAGCAGAGGACCTTCAAAACACTGCAACCTGGGGCCT 180
Db 2032 AATTCGAGATCGAATACAACTGATAGAGCAGAGGAGCTTCGAAACACTGGACCTGGGGCCT 2091

Qy 181 CCTCAGCCAAATGGATGCCCTGGACTCTCCCTTTCTTAGGACCTCTAGAGCTATATAATT 240
Db 2092 CCTCAGCCAAATGGATGCCCTGGATTCTCCCTTCTTAGGACCTCTAGAGCTATATAATT 2151

Qy 241 TTTTACTCTCTTTGGACCTCTGATCTTCAACTTCTTGTAGTATGTTTCTTCCAGAAT 300
Db 2152 GCTACTCTCTCTTTGGACCTCTGATCTTAACTCTCTTTAACTTTGTCTCTTCCAGAAT 2211

Qy 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAAACCCAGATGCAGTCCATGACTAA 360
Db 2212 CGAAGCTGTAAAGCTA-----CAATGGAGCCCAAGATGCAGTCCCAAGACTAA 2259

Qy 361 AATCTACCTGGACCCCTGGACCGGCTCTGATAGATATGCTCTGATGTTAATGACATTGA 420
Db 2260 GATCTACCCAGACCCCTGGACCGGCTCTGATAGCCACGATCTGATGTTAATGACATCA 2319

Qy 421 AGTCACCCCTCCCGAGGAAATCTCAACTGCACAAACCCCTACTACACTTCCAAATTCAGTAGG 480
Db 2320 AGTCACCCCTCTCTGAGGAAATCTCAGCTGCACAAACCTCTACTACGCCCAATTCAGCAGG 2379

Qy 481 AAGCAGTTAGAGCAGTTGTGAGCAACCTCCCAACAGTACTTTGGGTTTTTCTGTTGAGA 540
Db 2380 AAGCAGTTAGAGCGGTGCTGGGCCAACCTCCCAACAGCAGCTTAGGTTTTTCTGTTGAGA 2439

Qy 541 GGGTGGAGCTGAGAGCAGGACTAGCTGGATTCTCTAGGCTGACTAAGATCCCNAGCCT 600
Db 2440 TGGGGGACTGAGAGCAGGACTAGCTGGATTCTCTAGGCTGACTAAGATCCCNAGCCT 2499

Qy 601 ANCTGGGAAGTGCACCGCATCTCTTTAAACATGGGCTTGCAACTTAGCTCACACCCG 660
Db 2500 AGCTGGGAAGTGCACCATCCACATCCACCTTTAAACACGGGGCTTGCAACTTAGCTCACACCTG 2559

Qy 661 ACCAATCAGAGCTCACTAAATGCTTAATCAGGCAAAAACAGAGGTAAAGCAATAGCC 720
Db 2560 ACCAATCAGAGCTCACTAAATGCTTAATGAGCAAGCAGAGGTAAAGCAATAGCC 2619

Qy 721 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAATCAGGCA 780
Db 2620 AATCATCTATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAATCAGGCA 2679

Qy 781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTTTTTCTGTTTCTCCCTC 817
Db 2680 TTCAAGCCAGCAACAGCAACCCCTTTGGGTTTTTCTGTTTCTCCCTC 2716

RESULT 7
BX337769/c
LOCUS BX337769
DEFINITION BX337769 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI051YM13 3-PRIME, mRNA sequence.
ACCESSION BX337769
VERSION BX337769.2 GI:46272079
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 998)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30337641.

CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?S=CS0D1051AG07NP1&c=4215.r.

FEATURES
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1051YM13"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 50.4%; Score 669.4; DB 5; Length 998;
Best Local Similarity 89.2%; Pred. No. 2e-191;
Matches 754; Conservative 2; Mismatches 74; Indels 15; Gaps 3;

QY 1 TCAAAATCGAAGAGCTTTAGACTTCTGCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 60
DB 835 TCAAAATCGAAGAGCTTTAGACTTCTGCTAAACCGCTGAAAGAGGGGAACTGTTATTTT 776
QY 61 AGGGAAGAGTCTGTAGTATGTTAATCAATCTGGAATCATTACTGAGAAGTTAAAGA 120
DB 775 AGGGAAGAGTCTGTATATGTTAATCAATCCGGAATCGTCACTGAGAAGTTAAAGA 716
QY 121 AATTTGAGATCGAATATAATAGTACGACAGAGACCTTCAAAACATGCAACCTCGGGGCT 180
DB 715 AATTCGAGATCGAATACACGATAGACGAGGAGCTTCGAAACACTGGACCTGGGGCT 656
QY 181 CCTCAGCAATGATGCTCCCTGACTCTCCCTCTTAGGACCTCTAGCAGCTAATAATT 240
DB 655 CCTCAGCAATGATGCTCCCTGACTCTCCCTCTTAGGACCTCTAGCAGCTAATAATT 596
QY 241 TTACTCTCTTTGGACCTGATCTTCAACTCTCTTTGTAAGTTTGTCTCTCCAGAT 300
DB 595 GCTACTCTCTTTGGACCTGATCTTCAACTCTCTTTGTAAGTTTGTCTCTCCAGAT 536
QY 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAATGGAACCCCAAGATGCACTCACTAA 360
DB 535 CGAAGCTGTAAGCTA-----CAAAATGGAACCCCAAGATGCACTCACTAA 488
QY 361 AATCTACCTGGACCCCTGGACCGGCTGCTAGACTATCTCTGATGTTAATGACATTGA 420
DB 487 GATCTACCGGACCCCTGGACCGGCTGCTAGACTATCTCTGATGTTAATGACATTGA 428
QY 421 A-GTCAACCCCTCCGAGGAATCTCAACTGCAACCCCTACTACTACACTCCAATTCACTAG 479
DB 427 ACGGCAACCCCTCTGAGGAATCTCAGCTGCAACCTCTACTAGCCCAATTCAGCAG 368
QY 480 GAAGCAGTTAGACGATTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTGAG 539
DB 367 GAAGCAGTTAGACGATTTGTCAGCAACCTCCCAACAGTACTTGGGTTTCTCTGTGAG 308

QY 540 AGGCTGAGCTAGAGACAGACTAGCTGGATTTCTAGGCTGACTAAGATCCCAAGCC 599
DB 307 ATGGGGAGCTGAGAGACAGAGACTAGCTGGATTTCTAGGCTGACTAAGATCCCAAGCC 248
QY 600 TANTCGGAAGGTGACCGCATCCATCTTTAAACATGGGGTTTCAACTTTAGCTCACACC 659
DB 247 TAGCTGGGAAGGTGACCAATCCACCTTTAAACACGGGGCTTCAAACTTGTCTCACACT 188
QY 660 GACCAATCAGAGAGCTCACTAAATCTATCATGAGCAAAACAGGAGGTAAAGCAATAGC 719
DB 187 GACCAATCAGAGAGCTCACTAAATCTATCATGAGCAAAACAGGAGGTAAAGCAATAGC 128
QY 720 CAATCATCTATTCCCTGAGAGCAGCGGAGGAGGATTTGGGATATAAACTCAGGC 779
DB 127 CAATCATCTATTGCTGAGAGCAGCAGGAGGAGCAATGATCGGGATATAAAACCAAGT 68
QY 780 ATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTTCCCATTT--GTATGGGAGCTTGT 837
DB 67 CTTGAGCGGCAACGGCAACCCCTTTGGGTCCCNCCCTTTGTANTGGAGGTCTGT 8
QY 838 TTTC A 842
DB 7 TTTC A 3

RESULT 8
BX378303/c
LOCUS BX378303 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1018Y07 3-PRIME, mRNA sequence.
ACCESSION BX378303
VERSION BX378303.2 GI:46557492
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1058)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30439129.

CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?S=CS0D1018CD04NP1&c=4215.r.

FEATURES
Location/Qualifiers
1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1018Y07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 49.8%; Score 661.2; DB 5; Length 1058;
Best Local Similarity 86.5%; Pred. No. 6.2e-189;
Matches 727; Conservative 9; Mismatches 92; Indels 12; Gaps 1;
QY 1 TCAAAATCGAAGAGCTTTAGACTTCTGCTAAACCGCAAGAGGGGAACTGTTATTTT 60

```
Db      832 TCMAATCGNAGAGCTTTAGACTTGTACCGCTGAAGAGGKKKAWCYTTTTTTTTT 773
Qy      61  AGGGGAAGAAATGCTGTAGTATGTTAATCATCTGGAATCATTTACTGAGAAAGTTAAAGA 120
Db      772 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCKKAWTKTCACTGAGAAAGTTAAAGA 713
Qy      121 AATTTGAGATCGAATATATGATAGACAGAGGACCTTCAAAACACATGCACCTTGGGGCCT 180
Db      712 AATTCGAGATCGAATATACAGTAGAGAGAGGAGCTTCGAAACACATGAGACCTTGGGGCCT 653
Qy      181 CCTCAGCAATGATGCTTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240
Db      652 CCTCAGCAATGATGCTTGGACTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 593
Qy      241 TTTTACTCTCTTTGGACCTGTATCTTCAATCTCTTGTAAAGTTTGTCTCTCCAGAA 300
Db      592 GCTACTCTCTTTGGACCTGTATCTTAACTCTCTTGTAACTTTGTCTCTTCCAGAA 533
Qy      301 TGAAGCTGTAAGCTACAATAGTTCTTCAATGGAACCCAGATGCAGTCCATGACTAA 360
Db      532 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGAGTCCAAAGACTAA 485
Qy      361 AATCTACCGTGGACCTCTGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 420
Db      484 GATCTACCGCAGACCTCTGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCA 425
Qy      421 AGTCACCTCTCCGAGAGAAATCTCAACTGACAAACCCCTACTACTACCTCCAAATTCAGTAGG 480
Db      424 AGGCACCTCTCTGAGGAAATCTCAGCTGCACAACTCTACTACGCCCCCAATTCAGCAGG 365
Qy      481 AAGCAGTTAGACAGTTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCCTGTTGAGA 540
Db      364 AAGCAGTTAGACAGTTGTGAGCCAACTCCCAACAGTACTTGGGTTTTCCTGTTGAGA 305
Qy      541 GGGTGGACTCAGAGACAGGACTAGCTGATTTCTTAGGCTGACTTAAGAAATCCCNAAAGCCT 600
Db      304 TGGGGNACTGAGAGACAGGACTAGCTGATTTCTYTAGGCTGACTAAGAAATCCCTAAGCCT 245
Qy      601 ANCTGGGAAGTGACCGCATCCATCTTTTAAACATGGGGCTTGGCACTTAGCTACACCCG 660
Db      244 AGCTGGGAAGTGACCATCCACTTTAAACACGCGGCTTGAACCTTAGCTACACCTG 185
Qy      661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGCAGCAAAACAGAGGTAAGCAATAGCC 720
Db      184 ACCAATCAGAGAGCTCACTAAATGCTAATCAGCAGCAAAACAGAGGTAAGCAATAGCC 125
Qy      721 AATCATCTATTGCTGAGAGACAGCGGGAAGGACAGGATTTGGGATATATAAATCAGGCA 780
Db      124 AATCATCTATTGCTGAGAGACAGCGGGAAGGACAGGATTTGGGATATATAAATCAGGCA 65
Qy      781 TTCAAGCCAGCAACAGCAACCCCTTTGGGTCCTCCCTCCCATTTGATGGGAGCTCTGTTTT 840
Db      64  TTCGAGCGGCAACGGCAACCCNNTNTGNTNCCNTGTTGTTGTTGTTGTTGTTGTT 5
```

```
RESULT 9
BX439636/c
LOCUS
DEFINITION BX439636 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012Y24
3-PRIME, mRNA sequence.
ACCESSION BX439636
VERSION BX439636.2 GI:47000005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1019)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
On May 15, 2003 this sequence version replaced gi:30771765.
```

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?cs=CS0DE012DE12NP16c=4215.r.

Location/Qualifiers

FEATURES

source

1..1019

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DE012Y24"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed

with a NotI-oligo(dT) primer. Five prime end enriched,

double-strand cDNA was digested with NotI and cloned into

the NotI and EcoRV sites of the pCMVSPORT 6 vector.

Library was not normalized."

ORIGIN

Query Match 49.1%; Score 652.2; DB 5; Length 1019;

Best Local Similarity 85.5%; Pred. No. 3.3e-186; Indels 13; Gaps 2;

Matches 725; Conservative 19; Mismatches 91;

Qy 1 TCMAATCGAAGAGCTTTAGACTTGTCTAACCCGCAAAAGAGGGGGAACCTGTTATTTTT 60

Db 835 TCMAATCGAAGAGCTTTAGACTTGTCTAACCCGCTGAAGAGGGGGAACCTGTTATTTTT 776

Qy 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCATTTACTGAGAAAGTTAAAGA 120

Db 775 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCGTCACTGAGAAAGTTAAAGA 716

Qy 121 AATTTGAGATCGAATATATGATAGACAGAGGACCTTCAAAACACCTGCACCTTGGGGCCT 180

Db 715 AATTCGAGATCGAATATACACGCTAGACAGAGGAGCTTCGAAACACCTGGACCTTGGGGCCT 656

Qy 181 CCTCAGCAATGATGCTTGGACCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 240

Db 655 CCTCAGCAATGATGCTTGGACCTCTCCCTCTTTAGGACCTCTAGCAGCTATAATATT 596

Qy 241 TTTTACTCTCTTTGGACCTGTATCTTCAACTCTTCTTGTAAAGTTTGTCTTCCAGAA 300

Db 595 GCTACTCTCTTTGGACCTGTATCTTAAACCTCTTGTAACTTTGTCTTCTTCCAGAA 536

Qy 301 TGAAGCTGTAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAGTCCATGACTAA 360

Db 535 CGAAGCTGTAAACTA-----CAATGGAGCCCAAGATGCAGTCCAGACTAA 488

Qy 361 AATCTACCGTGGACCTCTGGACCGGCTGCTAGACTATGCTCTGATGTTAATGACATGA 420

Db 487 GATCTACCGCAGACCTCTGGACCGGCTGCTAGCCACGATCTGATGTTAATGACATCAA 428

Qy 421 AGTCACCTCTCCGAGGAAATCTCAACTGACAAACCCCTACTACTACCTCAATTCAGTAGG 480

Db 427 AGGCACCTCTCTGAGGAAATCTCAGCTGACAAACCTATATACGCCCCCAATTCAGCAGG 368

Qy 481 AAGCAGTTAGACAGTTGTGAGCCAACTCCCAACAGACTACTTGGGTTTTCCTGTTGAGA 540

Db 367 AAGCAGTTAGACAGTTGTGAGCCAACTCCCAACAGACTACTTGGGTTTTCCTGTTGAGA 308

Qy 541 GCGTGGACTGAGAGACAGGACTAGCTGATTTTCTTAGGCTGACTTAAGAAATCCCNAAAGCCT 600

Db 307 TGGGGNACTGAGAGACAGGACTAGCTGGATYCCYAGGCTGACTAAGAAATCCCTAAGCCT 248

Qy 601 ANCTGGGAAGGTGACCGCATCTTAAACATGTTGGGCTTGGCAATTAGCTACACCCG 660

Db 247 AGCTGGGAAGTGACCAATCCACCTTTAAACACGGGCTTGCAACTAGCTCACACTG 188
 Qy 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAAACGAGGTAAAGCAATAGCC 720
 Db 187 ACCAATCAGAGAGCTCACTAAATGCTAATCAGGCAAAAAACGAGGTAAAGCAATAGCC 128
 Qy 721 AATCATCTATTGCTTGGAGACACGCGGAGGACAGGATGGGATATAAATCAGGCA 780
 Db 127 RATTATTTATTGCTTGGAGACACGAGGAGGACATATCGGATATAAATCAGGCA 68
 Qy 781 TTCAAGCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATGATGGAGCTCTGTTTT 840
 Db 67 TTCGAGCGGCAACAGCAACCCCTTTGGGTCCCTCCCAATGATGGAGCTCTGTTTT 9
 Qy 841 CACTCTAT 848
 Db 8 NTTTATTT 1

RESULT 10
 BX380176/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 966)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 8, 2003 this sequence version replaced gi:30460243.
 Contact: Genoscope
 Genoscope
 Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 4215.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CSODI044BF03NP1&c=4215.r.
 Location/Qualifiers

FEATURES
 source
 1..966
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI044YK06"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 45.9%; Score 610; DB 5; Length 966;
 Best Local Similarity 82.5%; Pred. No. 2,2e-173;
 Matches 693; Conservative 24; Mismatches 110; Indels 13; Gaps 2;
 Qy 1 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTATTTT 60
 Db 830 TCAAAATCGAAGAGCTTTAGACTTGTCTAAACCGCAAAAGAGGGGAACTGTTATTTT 771
 Qy 61 AGGGGAAGATGCTGTAGTATGTTAATCATCTGAATCATCTAGCAAGAGTTAAAGA 120
 Db 770 AGGGGAAGATGCTGTATTTATTTATTAATCAATCCGAATCGTCACTGAGAAAGTTAAGA 711

Qy 121 AATTGAGATCGAATTAATGTAGAGCAGAGGACCTTCAAAAACATGCACTGCGGSCCT 180
 Db 710 AATTGAGATCGAATTAATGTAGAGCAGAGGACCTTCAAAAACATGCGGSCCT 651
 Qy 181 CCTCAGCCAAATGATGCGGCTGCTCCCTTTCTTAGGACCTCTAGCAGCTATATATT 240
 Db 650 CCTCAGCCAAATGATGCGGCTGCTCCCTTTCTTAGGACCTCTAGCAGCTATATATT 591
 Qy 241 TTTACTCTCTTTTGGACCCCTGATCTTCAACTTCTTTAAAGTTTGTCTCTTCCAGAA 300
 Db 590 GCTACTCTCTTTTGGACCCCTGATCTTAACTTCTTTGTTAACTTTGTCTCTTCCAGAA 531
 Qy 301 TGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAATGATGATA 360
 Db 530 CGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATGGAACCCAGATGCAATGATGATA 483
 Qy 361 AATCTACCGTGGACCCCTGAGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 420
 Db 482 GDTCTACCGGCGCCCTGAGACCGGCTGTAGACTATGCTCTGATGTTAATGACATTGA 423
 Qy 421 AGTCACCCCTCCGAGGAAATCTCAATGCAACACCCCTACTACACTCCCAATTCAGTAGG 480
 Db 422 AGGCACCCCTCCGAGGAAATCTCAATGCAACACCCCTACTACACTCCCAATTCAGTAGG 363
 Qy 481 AAGCAGTTAGAGCAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTTGTGAGA 540
 Db 362 AAGCAGTTAGAGCAGTTGTTCAGCAACCTCCCAACAGTACTTGGGTTTCTTGTGAGA 303
 Qy 541 GGGTGGACTAGAGACAGGACTAGTGGATTTCTAGGCTGACTAAGCAATCCCAAGCCT 600
 Db 302 TGGGGGACTAGAGACAGGACTAGTGGATTTCTAGGCTGACTAAGCAATCCCAAGCCT 243
 Qy 601 ANCTGGGAAGTACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 660
 Db 242 RGCTGGGAAGTACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCG 183
 Qy 661 ACCAATCAGAGAGCTCACTAAATGCTAATCAGCAAAAAACAGAGGTAAAGCAATAGC- 719
 Db 182 CCCAATCAGAGAGCTCACTAAATGCTAATCAGCAAAAAACAGAGGTAAAGCAATAGC- 123
 Qy 720 CAATCATCTATTGCTTGGAGCAGGACGAGGAGGACAGGATTCGGATATATAAATCTCAGGC 779
 Db 122 CMSVMTCTATTGCTTGGAGCAGGACGAGGAGGACAGGATTCGGATATATAAATCTCAGGC 63
 Qy 780 ATTCAAGCCAGCAACAGCAACCCCTTTGGGTCCCTCCCAATGATGAGGAGCTCTGTTT 839
 Db 62 CTTCGCGCGCCCGGCGCCCTTTGGGTCCCTCCCAATGATGAGGAGCTCTGTTT 3

RESULT 11
 BX357208/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 758)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 5, 2003 this sequence version replaced gi:30376125.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned

1..758
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI022YJ18 3-PRIME, mRNA sequence."
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4215.r

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?s=CS0DI0222DE09NP1&c=4215.r>.

FEATURES
SOURCE

```

1. . 738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="CS01022YJ18"
/tissue_types="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NciI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	45.8%	Score 508.2	DB 5	Length 758
Best Local Similarity	88.1%	Pred. No. 7.2e-173		
Matches	665	Conservative	9	Mismatches 69
			Indels	12
			Gaps	1
Qy	75	GTTAGTATGTTAATCAATCTGGATCATCTACGTAGAAAGTTAAAGAAATTTGAGATCGAA	134	
Db	758	GTTATTATGTTAATCAATTCGGATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAA	699	
Qy	135	TATAATGTAGACGACGAGACCTTCAAAACACTGCACCTCGGGGCTCTCTCAGCCAAATGGA	194	
Db	698	TACAACTGTAGACGACGAGAGCTTCGAAACACTGCAGCCCTGGGGCTCTCTCAGCCAAATGGA	639	
Qy	195	TGCCCTGGACTCTCCCTTCTTAGGACCTCTAGCAGCTATATAATTTTTACTCCTCTTTG	254	
Db	638	TGCCCTGGATTCTCCCTTCTTAGGACCTCTAGCAGCTATATAATTTGCTACTCTCTTTG	579	
Qy	255	GACCTGTATCTTCAACTTCCTTGTAAAGTTGTCTCTTCCAGAAATTTGAAGCTGTAAAGC	314	
Db	578	GACCTGTATCTTAAACCTCTTGTAACTTTGTCTCTTCAGAAATTCGAAGCTGTAAAGC	519	
Qy	315	TACAAATAGTCTTCAAAATGAAACCCAGATGCACTGCATGACTTAAATCTACCTGTGAC	374	
Db	518	TA-----CAAAATGGAGCCCAAGATGCACTGCAGTCCAAGCTTAAGATCTACCGCAGC	471	
Qy	375	CCCTGGACCGGCTGTAGACTATGCTCTGTATGTTAATGACATTCGAAGTCAACCTCCCG	434	
Db	470	CCCTGGACCGGCTGTAGCCACGATCTGATGTTAATGACATCAAGAGCAACCTCTCTG	411	
Qy	435	AGGAAATCTCAACTGCACAAACCCCTACTACACTCCAAATTCAGTAGGAAAGCAGTTAGACA	494	
Db	410	AGGAAATCTCAGCTGCACAACTCTACTACGCCCAATTCAGCAGAAAGCAGTTAGAGG	351	
Qy	495	GTTGTGACGCAACCTCCCAACAGTACTTGGGTTTTCTGTGTAGAGGGTGGACTGAGAG	554	
Db	350	GTGTCGGCCAACTCCCAACAGCAGACTTAGTGTCTGTGTAGATGGGGACTGAGAG	291	
Qy	555	ACAGGACTAGCTGGATTTCTTAGCTGACTAAGAAATCCCNAAAGCTTACTGGGAAGTTGA	614	
Db	290	ACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCTTAAGATCCCTTAAGCCTAG	231	
Qy	615	CGGATCCATCTTTAAACATATGGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGC	674	
Db	230	CCACATCCACCTTTAAACACGGGGCTTGCACTTAGCTCACACCTGACCAATCAGAGAGC	171	
Qy	675	TCACCTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCGCAATCATCTATTGCC	734	
Db	170	TCACCTAAATGCTAATTAGGCAAAAGCAGGAGGTAAAGAAATTAGCCATCATCTATTGCC	111	
Qy	735	TGAGAGCAGCGGGGAAGGACAGGATTTGGGATATAAACTCAGGCAATTCAGGCCAGCAAC	794	
Db	110	YSAGAGCAGCAGGAGGAGCAATGATCGGATATAAACCCMMWSYCCCCCGCCGCGCCAC	51	
Qy	795	AGCAACCCCTTTGGGTCCCCTCCCATTTGATGGG	829	

50 GGCMACCCCCYCTGGGTCCCTCCCTTGTATGGG 16

RESULT 12
BI087886

BI087886 771 bp mRNA linear EST 20-JUN-2000
LOCUS 62852690F1 NIH MGC 10 Homo sapiens cDNA clone IMAGE:493894 5',
DEFINITION mRNA sequence.
ACCESSION BI087886
VERSION BI087886
KEYWORDS mRNA sequence.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 771)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

FEATURES

```

1. 771
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/ciOne="IMAGE:4993894"
/cell_line="MGC36"
/lab_host="DH10B"
/clone_lib="NIH_MGC_10"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NO1; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 kb. Library prepared by Life Technologies."

```

ORIGIN

Query Match	44.7%;	Score 593.4;	DB 4;	Length 771;
Best Local Similarity	89.2%;	Pred. No. 2.3e-168;		
Matches 688;	Conservative 0;	Mismatches 68;	Indels 15;	Gaps 4;
QY	107	GAGAAAGTTAAAGAAATTTGAGATCGAATATAATGTAGAGCAGAGGACCTTTCAAAACACT	166	
DB	1	GAGAAAGTTAAAGAAATTCGAGATCGAATACAACTGAGCAGAGGACCTTCGAAACACT	60	
QY	167	GCACCTCGGGCGCTCCTCAGCCAATGGATGCCCTGGACTCTCCCTCTCTTAGGACCTCTA	226	
DB	61	GGACCTCGGGCGCTCCTCAGCCAATGGATGCCCTGGATTCTCCCTCTCTTAGGACCTCTA	120	
QY	227	GCAGCTATAATATTTTTACTCCTCTTTGGACCCCTGTATCTTTCAACTCCTCTGTTAGTTT	286	
DB	121	GCAGCTATAATATTTGCTACTCCTCTTTGGACCCCTGTATCTTTAACTCCTCTGTTAACTTT	180	
QY	287	GTCTCTTCAGAAATTTGAAGCTGTAAAGCTACAAATAGTTCCTTCAAATCGMAACCCAGATG	346	
DB	181	GTCTCTTCAGAAATCGAAGCTGTAAACCTA-----CAAATGGAGCCCAAGATG	228	
QY	347	CAGTCCATGACTTAAATCTACCGTGGACCCCTGGACCGGCTGTGTAGACTATGCTCTGAT	406	
DB	229	CAGTCCAGACTAAGATCTACGGACAGCCCTGGACCGGCTGTGTAGCCAGATCTGAT	288	
QY	407	GTTAATGACATTTGAAGTCACCCCTCCGAGGAAATCTCAAATGTCACAAACCCCTACTACAC	466	
DB	298	GTTAATGACATCARGGTCACCCCTCTCAGAGAAATCTCAGCTGCACAACTCTACTAGGC	348	

QY	467	TCCAAATCAGTAGGAGCAGTGTAGAGCAGTGTGTGACGCCAAGCTCCCAACAGTACTTGGG	526
Db	349		407
QY	527	TTTTCCTCT-TCAGAGGCTGACCTAGAGACAGAGCTAGCTGGATTTCTTAGGCTGACTA	585
Db	408		467
QY	586	AGAAATCCCAAGCCTTANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAA	645
Db	468		527
QY	646	CTTAGCTACACCCAGCCATCAGAGAGCTACTAAATGCTTAATCAGCGCAAAACACGGA	705
Db	528	CTTAGCTACACCTGACCAATCAGAGAGCTACTAAATGCTTAATGCGGAAAGACAGGA	587
QY	706	GGTA-AAGCAATAGCAATCATCTATTGCTGAGAGCACAGCGGGAAGACAAAGGATTGG	764
Db	588		647
QY	765	GATATAAATCTAGGATTCAGCCAGCAACAGCAACCCCTTTGGTCCCTTCCCATTTGT	824
Db	648	GATATTAAACCAAGCTTCGAGCGCGCAAGCGCAACCCCTTTGGTCCCTTCCCTTTGT	707
QY	825	ATGGGAGCTCTGTTTTCATCTCTATTTCATCTATTAAATCATGCAACTGCA	875
Db	708	ATGGGAGCTCTGTATTTCATCTATTTCATCTATTAAATCTTTGCAACTGCA	758
RESULT 13			
LOCUS	BC035153		
DEFINITION	Homo sapiens hypothetical protein FLJ22313, mRNA (cdna clone IMAGE:5265109), with apparent retained intron.		
ACCESSION	BC035153		
VERSION	BC035153.1	GI:23272890	
KEYWORDS	HTC.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 4204) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Schetz,T.E., Brongne,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Carninci,P., Prange,C., Raha,S.S., Bosak,S.A., McSwan,P.J., Abramson,R.D., Mullany,S.J., Garavito,R.M., Hume,D.A., Richardson,J., Mew,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Green,E.D., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Schmutz,J., Myers,R.M., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Skalska,U., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16899-16903	(2002)
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 4204) Strausberg,R.		
AUTHORS	Direct Submission		
TITLE	Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
JOURNAL			

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT	Contact: MGC help desk Email: cgabps-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amadan@systemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting			
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 73 Row: k Column: 22 This clone has the following problem: retained intron.			
FEATURES	Location/Qualifiers			
source	1..4204 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5265109" /tissue_type="Brain, hippocampus" /clone_lib="NIH MGC 95" /lab_host="DH10B" /note="Vector: pBluescript"			
ORIGIN	Query Match 44.5%; Score 592; DB 3; Length 4204; Best Local Similarity 81.1%; Pred. No. le-167; Matches 755; Conservative 0; Mismatches 128; Indels 48; Gaps 4; 400 CTCTGATGTTAATGACATTGAAGTCACCCCTCCCGAGGAATCTCAACTGCACAAACCCCT 459 887 CACTGAATTAGTTGAGAGACAGGCACCCCTCAGGAGGAATCTCACTGCACAAACCCCT 946 460 ACTACACTCCAAATTCAGTAGGAAGCAGTTAGAGCAGTTGTCTGAGCCAACTCTCCCAACAGT 519 947 ACTGTGCCCCAATTCAGCAGGAAGCAGTTAGAGCGGTGGTCA---AACCTCCCAATAGC 1003 520 ACTTGGGTTTTCTGTTGAGAGGGTGGACTGAGAGACAGGACTAGCTGGATTTCTCTAGGC 579 1004 ACTTGGGTTTTCTGTTGAGAGGGGTACTGAGAGACGAGACTAGTTGATTTCTCTAGGC 1063 580 TGACTAAGATCCNAAAGCCTANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGCG 639 1064 TGGCTAAGAAATCCCTAAGCCTAGCTGGGAATTGACCACGTCACCTTTAAACAGGGGC 1123 640 TTGCAACTTAGCTCACACCCGACCAATC-----AGAGAGCTCACTAAATGCTTAAT 590 1124 TTGCAATTTAGCTCACACCCGACCAATCAGGTAGTAAGAGAGCTCACTAAATGCTTAAT 1183 691 CAGCAAAACAGAGGTTAAGCAATAGCCAAATCATCTATTGCTGAGAGACAGCGGGA 750 1184 TAGGGAACAAACAGAGGTTAAGAAAGTAGCCAAATCATCTATCGCTGAGAGACACACGGA 1243 751 AGGCAAGGATTTGGATATAAACTCAGGCATTCAGCCAGCAACAGCAACCCCTTTGGG 810 1244 GGGCAATGATCAGGATATAAACCCAGGCATTCAGCCAGGTGGCTACCCCTTTTGGG 1303 811 TCCCTCCCATTTGATGGGAGCTCTGTTTTCATCTCTATTTCATCTATTAAATCATGCAA 870 1304 TCCCTCCCTTTGATGGAAGCTCTGTTTTCATCTCTATTAAATCTTGCAAATTGCA----- 1358 871 CTGCACTCTTCTGGTCCGTTGTTTATGCGTCAAGCTGAGCTTTGTTGCGCATCCACC 930 1359 ---CACATTTTCTGGTACGTGTGTGTACAGCTCAAGCTGAGCTTTGCTCACCGTCCACC 1415 931 ACTGCTGTTTGCCACCGTCAACAGACCCGCTGCTGACTTCCATCCCTTTGGATCCAGCAGA 990 1416 ACTGCTGTCTGCGCTGTCTACAGACCCACAGCTGACTTCCATCCCTCTG----- 1464 			

QY 991 GTGTCCACTGTGCTCTCTGATCAGCGAGGTACCCATTGGCACTCCGATCAGGCTAAAGG 1050
Db 1465 -----GATCAGCGAGCGCCCATTTGCTCTCTGATCGGCTAAAGG 1507
QY 1051 CTTGCAATTTCTCTGATGCTTAAGTGTCTGGTTTCTCTTAATAGAACTGAACACTGG 1110
Db 1508 CTTGCCCTTTGTTCTCTGAGGCTTAAGTGCACGTTTCTCTTAATAGAACTGAACACTAG 1567
QY 1111 TCACCTGGTTCATGTTCTCTTCCATGACCCACGCTTCTTAATAGAGCTATAAAGCTCA 1170
Db 1568 TCACCTGGTTCACAGATTCTCTTCCATGACCCACGCTTCTTAATAGAGCTATAAAGCTG 1627
QY 1171 CGCATGGCCCAAGATTCCATCTCTTGGTATCTGTGAGGCCCAAGAACCCAGTCCAGAGA 1230
Db 1628 CACTTTGGCCCAAGATTCCATCTCTTGGTATCCATGAGGCCCAAGAACCCAGTCCAGAGA 1687
QY 1231 ANGTGAGGTTGCCACATTTGGGAAGTGGCCACATGTCATTTGTAGCGGCCCAACAC 1290
Db 1688 ACACAAGACTTGGCCACATCTCGAAGTGGCCGCTCACCATCTTGGAGCGACCTGCCAC 1747
QY 1291 CATCTTGGGAGCTGTGGGAGCAGGATCCCC 1321
Db 1748 CATCTTGGGAGCTGTGGGAGCAGGATCCCC 1778

RESULT 14

CN272394 719 bp mRNA linear EST 16-MAY-2004
LOCUS 170060005410 GRN_PNEU Homo sapiens cDNA 5', mRNA sequence.
CN272394
VERSION CN272394.1 GI:47288808
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 719)
AUTHORS Brandenberger R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R.,
Lebkowski, J. and Stanton, D. W.
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert length: 719 Std Error: 0.00.
Location/Qualifiers
1. 719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/mol_type="mRNA"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN_PNEU"
/notes="oligo dt primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

FEATURES
source

RESULT 15
BX459153/c
LOCUS BX459153
DEFINITION 3-PRIME, mRNA sequence.
BX459153
ACCESSION BX459153.2
VERSION BX459153.2
KEYWORDS GI:47051796
SOURCE EST.
ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 900)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31021096.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library

ORIGIN

Query Match 42.2%; Score 561; DB 7; Length 719;
Best Local Similarity 85.4%; Pred. No. 1.6e-158;
Matches 657; Conservative 0; Mismatches 62; Indels 50; Gaps 1;
QY 95 GGAATCATTTCTGAGAAAGTTAAAGAAATTTGAGATCGAATATATGATAGACGAGGAC 154
Db 1 GGAATCGTCACTGAGAAAGTTAAAGAAATTCGAGATCGAATATATGATAGACGAGGAG 60

QY 155 CTTCAAAACACATGCAACCCCTGGGGCCCTCTCAGCCAAATGATGCCCTGGACTCTCCCTTC 214
Db 61 CTTCAAAACACATGCAACCCCTGGGGCCCTCTCAGCCAAATGATGCCCTGGACTCTCCCTTC 120
QY 215 TTAGCACTCTAGCAGCTATATAATTTTACTCTCTTTGGACCTGTATCTTTCAACTTC 274
Db 121 TTAGCACTCTAGCAGCTATATAATTTTACTCTCTTTGGACCTGTATCTTTCAACTTC 180
QY 275 CTTGTTAAAGTTTGTCTCTCTCCAGAAATGAAGCTGTAAAGCTACAAATAGTTCTTCAAAATG 334
Db 181 CTTGTT-----AATG 190
QY 335 GAACCCAGATGTCAGTCATGATAAATCTACCTGGACCCCTGGACCGGCTGCTAGA 394
Db 191 GAGCCCAAGATGTCAGTCATGATAAATCTACCTGGACCCCTGGACCGGCTGCTAGC 250
QY 395 CTATCTCTGATGTTAATGACATTCAGATTCACCTCCCGAGGAAATCTCAACTGCACAA 454
Db 251 CCACGATCTGATGTTAATGACATTCAGATTCACCTCCCGAGGAAATCTCAACTGCACAA 310
QY 455 CCCCTACTACACTCCAAATTCAGTAGGAAGCAGTTAGAGCAGTTGTGAGCAACCTCCCCA 514
Db 311 CTTCTACTAGCCCAATTCAGTAGGAAGCAGTTAGAGCAGTTGTGAGCAACCTCCCCA 370
QY 515 ACAGTACTTGGGTTTTCTCTGTTGAGGGTGGAGTGGAGCAGAGTCTAGTGGATTTCC 574
Db 371 ACAGCATTAGGTTTTCTCTGTTGAGATGGGGACTGAGAGCAGAGTCTAGTGGATTTCC 430
QY 575 TAGGCTGACTAGAAATCCNAGCCCTTACTGGAAGGTGACCGCATCTTTAAACAT 634
Db 431 TAGGCTGACTAGAAATCCNAGCCCTTACTGGAAGGTGACCGCATCTTTAAACAT 490
QY 635 GGGGCTTGCAACTTAGCTCACACCCGACCAATCAGAGAGCTCACTAAAAATGCTAATCAGG 694
Db 491 GGGGCTTGCAACTTAGCTCACACCTGACCAATCAGAGAGCTCACTAAAAATGCTAATCAGG 550
QY 695 CAAAACAGAGGTAAAGCAATAGCAATCATCTATTGCTGAGAGCAGAGCGGGAAGGA 754
Db 551 CAAAACAGAGGTAAAGCAATAGCAATCATCTATTGCTGAGAGCAGAGCGGGAAGGA 610
QY 755 CAAAGATTGGGATATAAATCTCAGGATTCAGCCGCAACAGCAACCCCTTTGGGTCCC 814
Db 611 CAAATGATCGGATATAAACCAGTCTTTCAGCGCGCAACGGCAACCCCTTTGGGTCCC 670
QY 815 CTCCCATTTGATGGGAGCTCTGTTTTCACCTCTATTTCACCTATTATAAT 863
Db 671 CTCCCATTTGATGGGAGCTCTGTTTTCACCTCTATTATAAT 719

BX459153 900 bp mRNA linear EST 05-MAY-2004
BX459153 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013Y120
3-PRIME, mRNA sequence.
BX459153
BX459153.2 GI:47051796
EST.
Homo sapiens (human)

was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 4215.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna7s=CS0DE013BE1ONP1&c=4215.r.

FEATURES

Location/Qualifiers
1..900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013Y120"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 40.9%; Score 544; DB 5; Length 900;
Best Local Similarity 77.5%; Pred. No. 2.5e-153;
Matches 631; Conservative 49; Mismatches 120; Indels 14; Gaps 3;
QY 1 TCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCAAAGAGGGGAACTGTTATTTT 60
DB 802 TCAAAATCGAAGAGCTTTAGACTTCTTAACCGCCGTAAGAGGGGAACTGTTATTTT 743
QY 61 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCTGGAATCAATTAAGAGAGTTAAAGA 120
DB 742 AGGGGAAGAAATGCTGTAGTATGTTAATCAATCCGAAATGCTGAGAAAGTTAAAGA 683
QY 121 AATTTGAGATCGAATATAATATAGAGAGAGACCTTCAAAACATGACCTGGGGCT 180
DB 682 AATTCGAGATCGAATACAACGTAGAGCAGAGGAGTTTCAAAACATGACCTGGGGCT 623
QY 181 CCTCAGCAATGGATGCTGACCTCCCTCTTAGGACCTTAGAGCTTATATATT 240
DB 622 CCTCAGCAATGGATGCTGACCTCCCTCTTAGGACCTTAGAGCTTATATATT 563
QY 241 TTTACTCTCTTTGGACCTGTATCTTCAACTTCTTGTAAAGTTGTCTTCAGAAAT 300
DB 562 GCTACTCTCTTTSGACCTG-ATCTTTAACTCTCTT-ATACTTTGCTCTTCCAGAA 505
QY 301 TGAAGTGTAAAGCTACAATAGTTCTTCAATAGGAACCCAGATGCACTCACTAA 360
DB 504 CGAAGCTGTAACCA-AAAAAAAAATGGAGCCAAAGATGCAAGCCCAASACCA 457
QY 361 AATCTACCTGGACCTCGACCGCTCTAGACTATCTCTGATGTTAATGACATTGA 420
DB 456 SACTTACCGAGACCTCTGACCGCCGCTGACCCATCTGATGTCATTAATTAACCA 397
QY 421 AGTCACCTCTCCGAGGAATCTCAATCGCAACCCCTACTACTCAATTCAGTAGG 480
DB 396 ASGCACCCCCGAGGAAATCCAGCCGACAACTCTACTACGCCCAATTCAGCAG 337
QY 481 AAGCAGTTAGAGAGTTGTCAGCAACTCCCAACAGTCTGGGTTTCTGTTGAGA 540
DB 336 AAGCAGTTAGAGAGTTGTCAGCAACTCCCAACAGTCTGGGTTTCTGTTGAGA 277
QY 541 GGGTGGACTGAGAGACAGGACTAGCTGGATTTCTTAGGCTGACTAAGAAATCCCAAGCCT 600
DB 276 TGGGCACTTSGAASACAGSACTAGCTTSSATCCMCCGCCCAACCCCAACCT 217
QY 601 ANCTGGGAAGGTGACCGCATCTTTTAAACATGGGGCTTGCAATTAGCTCACACCCG 660
DB 216 MGCTGGGAAGGTGACCAACCCMCCGCCCAACAGCGGGCCSCCACTTCGCTCAVCCTG 157
QY 661 ACCAATCAGAGCTCACTTAATGCTTAATCAGGCAAAACAGAGGTAAGCAATAGCC 720
DB 156 SCCAATCAGAGCGGSGGAAGGCTTTCAGGCAAGAGAGGAGGAGGATGAGCC 97
QY 721 AATCATCTATTGCTGAGAGCAGCGGGGAAGGCAAGGATTGGGATATAAACTCAGGCA 780

DB 96 MVCCATSCATGCTCTGAGAGACRGCCAGGAGGACVATGCGGGAGRTAAASCCAAGTC 37
QY 781 TTCAAGCCAGCAACAGCAACCCCTTTGGTCCC 814
DB 36 BCCGCGCGGTACGCGCKCCCCCTKTGGTCCC 3
RESULT 16
BX953822/c
LOCUS
DEFINITION
BX953822
DEFINITION DKFZP781M064.r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
BX953822
DEFINITION DKFZP781M064.5, mRNA sequence.
BX953822.1
DEFINITION GI:43433374
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1 (Bases 1 to 1002)
Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
Mewes,H.W., Well,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
EST (Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No 5' sequence available.
This clone (DKFZP781M064) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..1002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP781M064"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
ORIGIN
Query Match 37.7%; Score 500.6; DB 5; Length 1002;
Best Local Similarity 87.6%; Pred. No. 4e-140;
Matches 641; Conservative 0; Mismatches 72; Indels 19; Gaps 8;
QY 549 TGAGAGACAGCAGCTAGCTGGATTTCTAGGCTGATCCNAGCCCTACTGGA 608
DB 922 TGAGAGACAGCAGCTAGCTGGATTTCTAGGCTGATCCNAGCCCTACTGGA 863
QY 609 AGGTGACCGCATCTCTTTAAACATGGGGTTCCAACTTACCTCACACCCGCAATCA 668
DB 862 A-GTGACCGCATCTCTTTAAACATGGGGTTCCAACTTACCTCACACCCGCAATCA 808
QY 669 GAGAGCTCACTTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCT 728
DB 807 -----CTCCTTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCAATCATCT 753
QY 729 ATTGCTGAGAGCAGCGGGAAGCAAGGATTTGGGATATAAACTCAGGCAATCAAGCC 788
DB 752 ATTGCTGAGAGCAGCG-CGAAGGACAAAGATCGGGATATAAAACAGGCAATCAAGCC 694
QY 789 AGCAACAGCAACCCCTTTGGGTCCTCCCTCTGATGGAGCTCTGTTTTCACTCTAT 848

```
|||||
693 AGCAATGGCAACCCCTTTGGTCCCTCCCTCGTATGGAGCTCTGTTTCACTCTAT 634
QY TTCACCTATTAATCATGCACTGCACTCTTCTGCTCGTGTGTTTATGCTCAAGCT 908
Db TTACACACTATTAATCTTGAACATGCACTCTTCTGCTCGTGTGTTTATGCTCGAGCT 574
QY GAGCTTTTTCGCGCATCCACCACTGCTGTT-TGCCACCGTCCACAGCCCGTGTGACT 967
Db GAGCTTTTCGCTACTGTCACCACTGCTGTTTTCGCGGTAGCAGACTCGTGTGACT 514
QY TCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCAGGTA-CCCAT 1026
Db TACATTTCTCCGATCCAGCAGGCTGCCACTGTCTCTGATCCAGCAGGACCCCTT 454
QY TGCACTCCCGATCAGGCTAAAGCTTGCCA-TTGTCTCTGATGCTAAGTCCCTGGCT 1085
Db TGCCGCTCCAGATCGGGCTAAAGCTTGCCATTTGTCTGCAAGCTAAGTCCCTGGGT 394
QY TTGCTCTAATAGACTGAACACTGCTGCTGCTTCCATGGTCTCTTCCATGACCCACG 1145
Db TTGCTCTAATAGACTGAACACTGCTGCTGCTTCCATGGTCTCTTCCATGACCCACG 334
QY GCTTCTAATAGACTGAACACTGCTGCTGCTTCCATGGTCTCTTCCATGGTCTCTG 1205
Db ACTTCTAATAGACTGAACACTGCTGCTGCTTCCATGGTCTCTTCCATGGTCTCTG 278
QY GAGGCAAGAACCCAGGTCAGAGANGTGGAGCTTGCACCAT-TTGGGAAGTGGCCCA 1264
Db GAGGCAAGAACCCAGGTCAGAGAACACAGAGCTTGCACCATCTTTGGGAAGTGGCCCA 218
QY CTGCCATTTGG 1276
Db CGACCATCTTTG 206
```

```
RESULT 17
AG113694/c
LOCUS Pan troglodytes DNA, clone: PTB-120G11.R, genomic survey sequence.
AG113694
ACCESSION AG113694
VERSION AG113694.1 GI:16734213
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 679)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..679
/organism="Pan troglodytes"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-120G11.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
```

ORIGIN

```
Query Match 36.2%; Score 480.8; DB 9; Length 679;
Best Local Similarity 86.7%; Pred. No. 3.7e-134;
Matches 552; Conservative 0; Mismatches 83; Indels 2; Gaps 2;

QY 673 GTCCTACTAAATGCTAATCAGGCAAAAACAGAGAGTAAAGCAATAGCAATCATCTATTG 732
Db 678 GCTCGTAAAGTCTAATAAGGCAAAAACAGAGAGTAAAGCAATAGCAATCATCTATTG 619
QY 733 CCGTGAAGACACGCGGAGGACAAAGGATTGGGATATAAACTCAGGCACTTCAAGCCAGCA 792
Db 618 CGTGAAGACACAGTGTAAAGGAAAAAGGTCGGGATATAGCCCGGCAATTCGAGCCGCA 559
QY 793 AC-AGCAACCCCTTTGGGTCCCTCCCATTTGTATGGAGCTCT-CTTTTCACTCTATTT 850
Db 558 AGGGGAGCCCTTTGGGTCCCTCCCTTTGTATGGAGCTCTGCTTTTCACTCTATTT 499
QY 851 CACTCTATTAAATCATGCAACTGCACTCTTCTGCTCGGTGTTTATAGCTCAAGCTGA 910
Db 498 CACTATATTAGATCTTTCAGAGTGCACCTCTCTGCTCGGTGTTTATAGCTCAAAACAGA 439
QY 911 GCTTTTGTTCGCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCTGCTGACTTC 970
Db 438 GCTTTTGTTCGCGTCCGCGCTGCTGTTTACCGCGCTCCAGACCCGCTCAGCTGCTTC 379
QY 971 ATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGGTTACCATTTGCC 1030
Db 378 ATCCCTTCGATCCGCGAGGCTGTGCGCTGTGCTACTGATCCAGCAGGCAACCATTTGCT 319
QY 1031 ACTCCGATCAGGCTAAAGGCTTGCATTTGCTCGATGGCTAAGTGCCTGGTTGCTC 1090
Db 318 GCTCCGATCGGCTAAAGGCTTGCATTTGCTCGATGGCTAAGTGCCTGGTTGCTC 259
QY 1091 CTAAATGAAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
Db 258 GTAAATGAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199
QY 1151 TAATAGAGTGTAACTCACTCAGCAGTGGCCCAAGATTCCATTTCTTGGTATCTGTGAGGC 1210
Db 198 TAATAGAGTGTAACTCACTCAGCAGTGGCCCAAGATTCCGTTCTTGGATCCGTGAGGC 139
QY 1211 CAAGAACCCAGCTCAGAGANGTGGCTTGCACCATTTGGGAAGTGGCCCACTGCCA 1270
Db 138 CAAGAACCCAGCTCAGAGAACACAGAGGCTTGCACCATCTTGGGAAGTGGCCCACTGCCA 79
QY 1271 TTTTGTAGGCGCCACCATCTTGGGAGCTGTGG 1307
Db 78 TTTTGAAGTGGCGCCCATCTTGGGAGCTGTGG 42
```

RESULT 18

BM918330/c

LOCUS

DEFINITION

5', mRNA sequence.

BM918330

BM918330.1 GI:19368709

EST.

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 999)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

BM918330 999 bp mRNA linear EST 12-MAR-2002
AGENCOURT_6708649 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5747240
5', mRNA sequence.

BM918330
BM918330.1 GI:19368709
EST.
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 999)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12773 row: b column: 09 High quality sequence stop: 685. Location/Qualifiers	
	1..999 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5747240" /lab_host="DH10B" /clone_lib="NIH_MGC_120" /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."	
FEATURES	source	
	RESULT 19 AG134524 LOCUS DEFINITION Pan troglodytes DNA, clone: PTB-147110.R, genomic survey sequence. ACCESSION AG134524 VERSION AG134524.1 GI:16664202 KEYWORDS GSS. SOURCE Pan troglodytes (chimpanzee) ORGANISM Pan troglodytes Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan. REFERENCE 1 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. TITLE BAC end sequences of Library PTB JOURNAL Unpublished REFERENCE 2 (bases 1 to 683) AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y. TITLE Direct Submission JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimps@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/ , Tel: 81-45-503-9111, Fax: 81-45-503-9170) COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors. PRIMERS Sequencing: M13Rev LIBRARY Vector : pKS145 R.Site 1 : SacI R.Site 2 : SacI. Location/Qualifiers 1..683 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" /clone="PTB-147110.R" /sex="male" /cell_type="lymphoblast" /clone_lib="PTB Chimpanzee Male BAC Library"	
ORIGIN	source	
	Query Match 35.8%; Score 475.2; DB 9; Length 683; Best Local Similarity 89.6%; Pred. No. 1.8e-132; Matches 554; Conservative 0; Mismatches 60; Indels 4; Gaps 4; QY 1134 CCATGACCCACGGCTTCTAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTC 1193 Db 287 CCAATGACCCACGGCTTCTCAATAGAGCTATAACACTCACCGCATGGCCCAAGATTCCATTC 228 QY 1194 CTTGGTATCTGTGAGGCCAAGAACCCAGGTGAGAGANGTGAAGCTTGGCCACCATTTGG 1253 Db 227 CTTGGAATCGTGGGGCCAAAGAACCCAGGTGAGAGAACGAGAGGCTTGTGCGCATCTTG 168 QY 1254 GAAGTGGCCCACTGCCATTTTGGTAGC 1280 Db 167 GAAGCAACCCGCCACCATGTTGGAGC 141 Query Match 36.2%; Score 480.8; DB 5; Length 999; Best Local Similarity 83.1%; Pred. No. 4.1e-134; Indels 16; Gaps 6; Matches 621; Conservative 0; Mismatches 110; Indels 16; Gaps 6; QY 549 TGAGAGACAGGACTAGCT-GGATTTTCTAGGCTGACTAAGAAATCCNAAGCCTTACTGGG 607 Db 886 TGAGAGACAGGACTAGCTGGGATTTCTAAGCGGACTAAGAAATCCCTAGGCTAGTTAGG 827 QY 608 AGGTGACCGCATCCATCTTTAA-CATGGGCTTGCACTTAGCTCACACCCGACCAAT 666 Db 826 AAGGTGACCGCATCCACCTTAAACCATGGGGCTTGCAACTTAGCTCACACCCGACCAAT 767 QY 667 C-----AGAGACTCACTAAATGCTAATCAGGC-AAAAACAGGAGGTAAGCAAT 716 Db 766 CAGGTAGTAAGAGAGCTCACTAAATGCCATCAGACAAAACAGGAGGTAAGAACT 707 QY 717 AGCCAATCATCTATTCCTGAGAGCACAGCGGAGGACAGAGATTGGATATAACTCA 776 Db 706 AACCAATCATCTATTGCCTGAGAGCACAGTGGGAGGACAAATGATCAGGATATAAGCCA 647 QY 777 GGCATTCAAGCCAGCAACAGCACCCCTTTGGGTCCTCCCATTTGATGGAGCTGTG 836 Db 646 GGCATTCCAGCCGGCAACGGCTACCTCTTTGGCTCCCTCCCTTTGATGGAGCTGTG 587 QY 837 TTTTCACTCTATTTCCTACTCTATTAAATCATGCAACT---GCATCTTCTGGTCCGTTT 893 Db 586 TTTTCACTCTATTAAATCTTGTAACTGCATACCTCTTCTGTCGCTGTTGTACCGGTGTT 527 QY 894 TTTATGGCTCAAGCTGAGCTTTTGTTCGCCATCCACACTGCTGTTTGGCCACGTCACAG 953 Db 526 GTTACCGCTCAAGCTTAAGCTTTGCTGCTGTTGTCACCACTGCTGTTTGGCGAGTCGAG 467 QY 954 ACCGCTGCTGACTTCCATCCCTTTGGATCCAGAGTGTCCACTGTGCTCTGATCCA 1013 Db 466 ACCGCGCGCTGACTTCCACCCCTCCGGATCTGGCAGGGTGTCTGCTGTCTCTGATCCA 407 QY 1014 GCGAGGTACCCATTGCCACTCCCGATCAGGCTAAAGGCTTGCATTTCTCGATGGCT 1073 Db 406 GCGAGGCGCCCACTGCTGCTCCGATCGGGCTGAAGGCTTGCCTGTTTCTTCCACGGC- 348 QY 1074 AAGTGCCTGGGTTTGCTCTAATAGAACTGAACCTGGTCACTGGTTCATGGTCTCTT 1133 Db 347 AAGTGCCTGGGTTGCTCTAATGAGATGAACACTAGTCTGCTGGGTTCTACGGGTTCTCTT 288	

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.
 PRIMERS
 Sequencing: -21M13
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI
 Location/Qualifiers
 1. 681
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-124K18.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"
 ORIGIN
 Query Match 35.6%; Score 473; DB 9; Length 681;
 Best Local Similarity 88.8%; Pred. No. 8.6e-132;
 Matches 538; Conservative 0; Mismatches 56; Indels 12; Gaps 2;
 671 GAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 730
 Db 86 GAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 145
 731 TCCTGAGAGCAGCGGAGGAGGAGGATGGGATATAAAGTCAAGGATCAAGCCAG 790
 Db 146 TCCTGAGAGCAGCGGAGGAGGAGGATGGGATATAAAGTCAAGGATCAAGCCAG 205
 791 CAACAGCAACCCCTTTGGGTCCTCCATGCTATGATGGAGGCTCTGTTTCACTATTT 850
 Db 206 CAACGCTACCTCTTTGGGTCCTCCATGCTATGATGGAGGCTCTGTTTCACTATTT 255
 851 CACTCTATTAATCATGCAACTGCA--CTCTTCTGGTCCGCTGTTTATGGCTCAAGCT 908
 Db 256 CACTCTATTAATCTTTGCAACTGCACTCTCTTCTGGTCCGCTGTTTATGGCTCAAGCT 315
 909 GAGCTTTTGTGCGCATCCACACTGCTGTTGCGCACCGTCAAGACCGCTGCTGACTT 968
 Db 316 GAGCTTTTGTGCGCATCCACACTGCTGTTGCGCGCCGCGCAGACCGCTGCTGACTT 375
 969 CCATCCCTTTGGATCCAGCAGAGTGTCACCTGCTCTGATCCAGGAGGTACCATTTG 1028
 Db 376 CCATCCCTCCAGATCTGGCAGGGGTGCCACTGCTGCTGATCCAGGAGGCGCCCATG 435
 1029 CCATCCCATCAGGCTAAGGCTTGGCATTTGTTCTGCTGATGCTTAAGTGGCTTTG 1088
 Db 436 CCCTCCCATGAGGGCTAAGGCTTGGCATTTGTTCTGCTGATGCTTAAGTGGCTTTG 495
 1089 TCCTAATAGACTGAACACTGGTCACTGGGTTCCATGGTTCTCTCCATGACCCAGGCT 1148
 Db 496 TCCTAATAGACTGAACACTGGTCACTGGGTTCCATGGTTCTCTCCATGACCCAGGCT 555
 1149 TCCTAATAGACTGAACACTGGTCACTGGGTTCCATGGTTCTCTCCATGACCCAGGCT 1208
 Db 556 TCCTAATAGACTGAACACTGGTCACTGGGTTCCATGGTTCTCTCCATGACCCAGGCT 615
 1209 GCCAAGAACCCAGGTCAGAGANGTGGGCTTGGCCACCATTTGGGAAGTGGCCCACTGC 1268
 Db 616 GCCAAGAACCCAGGTCAGAGANGTGGGCTTGGCCACCATTTGGGAAGTGGCCCACTGC 675
 1269 CATTTT 1274

Db 676 CATTTT 681
 |||||
 AG096341 710 bp DNA linear GSS 03-NOV-2001
 Pan troglodytes DNA, clone: PTB-097K23.F, genomic survey sequence.
 AG096341
 DEFINITION
 ACCESSION
 VERSION AG096341.1 GI:16716858
 KEYWORDS
 SOURCE GSS.
 ORGANISM
 Pan troglodytes (chimpanzee)
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB
 2
 Unpublished
 2 (bases 1 to 710)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.
 PRIMERS
 Sequencing: -21M13
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI
 Location/Qualifiers
 1. 710
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-097K23.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"
 ORIGIN
 Query Match 35.3%; Score 469.6; DB 9; Length 710;
 Best Local Similarity 85.1%; Pred. No. 9.4e-131;
 Matches 561; Conservative 0; Mismatches 65; Indels 33; Gaps 2;
 671 GAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 730
 Db 75 GAGCTCACTAAATGCTAATAGGCAAAACAGGAGGTAAAGCAATAGCCATCATCTAT 134
 731 TCCTGAGAGCAGCGGAGGAGGATGGGATATAAAGTCAAGGATCAAGCCAG 790
 Db 135 TCCTGAGAGCAGCGGAGGAGGATGGGATATAAAGTCAAGGATCAAGCCAG 194
 791 CAACAGCAACCCCTTTGGGTCCTCCATGCTATGGAGGCTCTGTTTCACTATTT 850
 Db 195 CAACGCAACCCCTTTGGGTCCTCCATGCTATGGAGGCTCTGTTTCACTATTT 244
 851 CACTCTATTAATCATGCAACTGCTCTTCTGGTCCGCTGTTTATGGCTCAAGCTGA 910
 Db 245 CACTCTATTAATCATGCAACTGCTCTTCTGGTCCGCTGTTTATGGCTCAAGCTGA 304
 911 GCTTTGTTGCGCATCCACCACTGCTGTTTGGCACCGTGCACAGACCCGCTGCTGCC 970
 Db 305 GCTTTGTTGCGCATCCACCACTGCTGTTTGGCACCGTGCACAGACCCGCTGCTGCC 364

Db 57 GAGAGCTCACTAAAAAGCTAATTAGGCAAAAAACAGGAGGTAAAGAAATAGCCAAATCATTT 116
Qy 729 ATTGCTGTGAGACACAGCGGAGGACAGGATTTGGGATATTAACCTCAGCATTCAAGCC 788
Db 117 ATTGCTGTGAGACACAGCGGAGGACAGGATTTGGGATATTAACCCAGGATTTGAGCC 176
Qy 789 AGCAACAGCAAAACCCCTTTGGGTCCCTCCATTTGATGGAGCTCTGTTTTCACTCTAT 848
Db 177 AGCAACGGGTACCTCTTTGGGTCCCTCCCTTTGATGGAGCTCTTTTTCACCTCTAT 236
Qy 849 TTCACCTATTAATAATCATGCAACTGCACTCTTCTGTCGCTGTTTTTATGCTCAAGCT 908
Db 237 TAAATCTTGCAACTGC-----GCTCTCTTCTGCTGCGTTTGTACAGCTCAAGCT 288
Qy 909 GAGCTTTTGTTCGCCATCCACCACTGCTGTTTGGCCACCGTCACAGACCCGCTGCTCACTT 968
Db 289 GAG-TTTTGTGCTGTGTCACCACTGCTGTTTCCGCGCTCGAGACCCGCTGCTGAGTT 347
Qy 969 CCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCTGATCCAGCGAGGTACCCATTG 1028
Db 348 CCATCCCTCGGATCCACAGGGTGTCCACTGTGCTCTGATCCAGCAAGCGCCCATG 407
Qy 1029 CCATCCGATCAGGCTAAAGGCTTGCATTTCTCTGCAATGCTAAGTCCCTGGGTTG 1088
Db 408 CCATCCGATCAGGCTAAAGGCTTGCATTTCTCTGCAATGCTAAGTCCCTGGGTTG 467
Qy 1089 TCCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148
Db 468 TCCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
Qy 1149 TCCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208
Db 528 TCCTAATAGACTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 587
Qy 1209 GCCAAGAACCCAGGTCAGAGANGTGAAGCTTGCACCAATTTGGGAAGTGGCCCACTGC 1268
Db 588 GCCAAGAACCCAGGTCAGAGANGTGAAGCTTGCACCAATTTGGGAAGTGGCCCACTGC 647
Qy 1269 CATTTTGGTAGC 1280
Db 648 CATCTTGGCGATC 659

RESULT 26
AG099717
LOCUS
DEFINITION
Pan troglodytes DNA, clone: PTB-102B07.F, genomic survey sequence.
AG099717
ACCESSION
AG099717.1 GI:16720234
VERSION
GSS.
KEYWORDS
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
REFERENCE
1
AUTHORS
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE
BAC end sequences of Library PTB
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 712)
AUTHORS
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE
Direct Submission
JOURNAL
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukuba, Ibaraki, 305-8565, Japan
(E-mail: chimpanzee@gsr.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/).
Tel: 81-45-503-9111 Fax: 81-45-503-9170
COMMENT
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13

FEATURES
source

LIBRARY
Vector : pKSI45
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1..712
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-102B07.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 34.2%; Score 455; DB 9; Length 712;
Best Local Similarity 86.2%; Pred. No. 2,6e-126; Indels 9; Gaps 2;
Matches 529; Conservative 0; Mismatches 76;
Qy 671 GAGCTCACTAAAATGCTAATCAGGCAAAAAACAGGAGGTAAAGCAATAGCCAAATCATCTAT 730
Db 107 GAGCTCACTAAAATGCTAATTA-GCAAAAAACAGGAGGTAAAGAAATAGCCAAATCATCTAT 165
Qy 731 TGCCTGAGAGCAGCGGGAAGGACAAAGATTGGGATATAAACTCAGGCAATTCAGGCCAG 790
Db 166 TGCCTGAGAGCAGCGGGAAGGACAAATGATCGGGATATAAAACCCAGGCAATTCAGGCCAG 225
Qy 791 CACAGCAACCCCTTTGGGTCCCTCCATTTGATGGGAGCTCTGTTTTCACTCTATTT 850
Db 226 CAAGGGCTACCTCTTTGGGTCCCTCCCTTTGATGGGAGCTCTGTTTTCACTCTATTT 284
Qy 851 CACTCTATTAATATCATGCAACTGCACTCTTCTGGTCCGCTGTTTTTATGGCTCAAGCTGA 910
Db 285 -----ATATCTCGCAACTGCACTCTCTCTGGCCTGTATTTGTTATGGCTCTAGCTGA 337
Qy 911 GCTTTTGTGCGCATCCACACTGCTGTTTGCACCGTCACAGACCCGCTGCTGACTTCC 970
Db 338 GCTTTTGTGAGCGTCCACCACTGCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 397
Qy 971 ATCCCTTTCGATCCAGCAGAGTGTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030
Db 398 ATCCCTTTCGATCCAGCAGAGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457
Qy 1031 ACTCCCGATCAGGCTAAAGCTTTGCCATTTCTCTGATGGCTTAAGTGCCTGGGTTGTC 1090
Db 458 ACTCCCGATTTGGCTTAAGACTTTGCCATTTGTTCCCGCATGGCTAAGTGCCTGCTGCTATC 517
Qy 1091 CTAATGAACCTGAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
Db 518 CTAATCGAGCTGAACACTAGTCACTGGGTTCCACAGTTCTCTTCCATGATCCATGGGCTTC 577
Qy 1151 TANTAGCTATTAACACTCACCAGCATGGCCCAAGATTCATTTCTTGGTATCTGTGAGGC 1210
Db 578 TANTAGCTATTAACACTTACTGATGGCCGAAGATCTCATTTCTTGGAACTGTGAGGC 637
Qy 1211 CAAGAACCCAGGTCAGAGAAAGTGGAGGCTTGGCCACCATTTGGGAAGTGGCCCACTGCCA 1270
Db 638 CAAGAACCCAGGTCAGAGAAAGTGGAGGCTTGGCCACCATTTGGGAAGTGGCCCACTGCCA 697
Qy 1271 TTTTGTGAGCGGC 1284
Db 698 TCTTGAAGAGGCC 711

RESULT 27
AG121669/c
LOCUS
DEFINITION
Pan troglodytes DNA, clone: PTB-130M15.F, genomic survey sequence.
AG121669
ACCESSION
AG121669.1 GI:16650834
VERSION
GSS.
KEYWORDS
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Toki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 689)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Toki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. 689
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-130M15.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
33.0%; Score 439.2; DB 9; Length 689;
Best Local Similarity 87.0%; Pred. No. 1.6e-121; Indels 10; Gaps 2;
Matches 507; Conservative 0; Mismatches 56;

ORIGIN
Query Match 33.0%; Score 438.8; DB 8; Length 921;
Best Local Similarity 83.8%; Pred. No. 2.4e-121; Indels 22; Gaps 4;
Matches 550; Conservative 0; Mismatches 84;

QY 257 CCTGTATCTTCAACTTCCTTGTAAAGTTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 316
DB 666 CCGTGGATCTTAAAGCTCTTGTAAAGCTTGTCTCTCCAGAAATGAAGCTGTAAAGCTA 607
QY 317 CAAATAGTTCTTCAATGAACCCGAGATGAGTCAATCAATCAATCAATCAATCAATCAAT 376
DB 606 TAATAGTTCTTCAATGAACCCGAGATGAGTCAATCAATCAATCAATCAATCAATCAAT 547
QY 377 CTGGACCGCTGCTAGACTGTCTGTATGTTAATGACATTTGAAGTCAATCAATCAATCAAT 436
DB 546 CTGGACCGCTGCTAGACTGTCTGTATGTTAATGACATTTGAAGTCAATCAATCAATCAAT 487
QY 437 GAAATCTCAATCTGAC- AACCCCTACTACATCTCAATCAATCAATCAATCAATCAATCAAT 495
DB 486 GAAATCTCAATCTGAC- AACCCCTACTACATCTCAATCAATCAATCAATCAATCAATCAAT 427
QY 496 TTGTGACCCACCTCCCAACAGTACTTGGGTTTCTGTTGAGGGGTGAGTCAATGAGAGA 555
DB 426 TCGTCGCGCAACCTCCCAACAGTACTTGGGTTTCTGTTGAGGGGTGAGTCAATGAGAGA 367
QY 556 CAGGACTAGTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGGGAAGTGCAC 615
DB 366 CAGGACTAGTGGATTTCTTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGGGAAGTGCAC 307
QY 616 CGCATCCATCTTTAAACATGGGGCTTGCACCTTAGCTCAACCCGACCAATC----- 667
DB 306 CACATTCACCTTTTAAACATGGGGCTTGCACCTTAGCTCAACCCGACCAATC----- 247
QY 668 -AGAGAGCTCAGTAAATGCTATCAGGCAAAACAGGAGGTAAAGCAATAGCCATCAT 726
DB 246 AAGAGGCTCAGTAAATGCTATCAGGCAAAACAGGAGGTAAAGCAATAGCCATCAT 187
QY 727 CTATTGCTGAGAGCAGCGGGAAGCAGAGGATTTGGGATATAAACTCAGGCAATTCAGG 786

186 TTATTGCTGAGAGTACAGCGGAGGACAAATGATCGGATATAAACTCATGCAATTCGAG 127
QY 787 CCAGCAACAGCAACCCCTTTGGTCCCTCCCTCCCAATGTTATGG 829
DB 126 CCACCAATGGCTACCTCTTTGGGTCCTCCCTCTTTGTTATGG 84

RESULT 28
BH149565/c
LOCUS
DEFINITION
ENTQ48FR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION
BH149565
VERSION
BH149565.1 GI:15310303
KEYWORDS
GSS.
SOURCE
Entamoeba histolytica
ORGANISM
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE
1. (bases 1 to 921)
Loftus B., Wang Z., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
HM1:IMSS sheared DNA library (2001)
Unpublished (2001)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjoftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 40
High quality sequence stop: 567.

FEATURES
source
1. 921
Location/Qualifiers
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

ORIGIN
Query Match 33.0%; Score 438.8; DB 8; Length 921;
Best Local Similarity 83.8%; Pred. No. 2.4e-121; Indels 22; Gaps 4;
Matches 550; Conservative 0; Mismatches 84;

QY 547 ACTGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGG 606
DB 651 AGTGAGACAGGACTAGCTGGATTTCCTAGGCTGACTAAGAAATCCCNAAAGCTTANCTGG 592
QY 607 GAAGTGACCGCATCCATCTTTAAACATGGGGCTTGCACCTAGCTACACCCGACCAAT 666
DB 591 GAAGTGACTGCATCCACCTATAAACACGGGGCTTGCACCTAGCTACACCCGACCAAT 532
QY 667 C-----AGAGAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATA 717

Db 531 CAGGTAGTAAGAGAGCTCACTAAACAGCTAATAGGCAAAACAGGAGGTAAGAAATA 472
 QY 718 GCCATCATCTATTGCTGAGAGCGCAGCGGAGACAGAGATGGGATATAAACTCAG 777
 Db 471 GCCAATCATCTATGCGCTGAGAGCACACCTGGGAGGACAAATGA-TAGGATATAAACCCAG 413
 QY 778 GCATTCAAGCCAGCAACAGCAACCCCTTTGGGTGCTCCCTCCCATCTGATGGAGCTCTGT 837
 Db 412 GCATTCAACAGCAGGAGGCTACCTCTTTGAGTCCCTCCCTTTGATGGAGCTCTGT 353
 QY 838 TTTCACTCTATTTCCTCTATTAAATCATGCAACTG-CACCTCTGCTCGTGGTCTTT 895
 Db 352 -----TTTCACTCTATTAAATCTTGCAACTGCACTCTCTGCTGCTGCTTTGT 303
 QY 896 TATGGCTCAAGCTGAGCTTTGTTCCGCATCCACCACTGCTGTTGCCACCGTCAAGAC 955
 Db 302 TAGGACTCGAGCTGAGCTTTCACTTTCGGCTCCACCACTGCTGTTGCCGCATCGAGAC 243
 QY 956 CGCTGCTGACTTCCATCCCTTTGGATCCAGCAGAGTGTCCACTGTGCTCCTGATCCAGC 1015
 Db 242 CGCGCTGACTTCCACCTCCAGATCTGGCAGGGTGTCTGCTGATCCTGATCCAGT 183
 QY 1016 GAGGTACCCATGCTCCACTCCGATCAGGCTAAAGCTTGGCATTGTTCTGCAATGCTTAA 1075
 Db 182 GAGGACCCATGCGCTCTGGATTGGGCTAAAGTCTTGTCACTTGTCTTTCACGCTAA 123
 QY 1076 GTGCTGGGTTGCTCTAATAGACTGAACACTGCTGCTGCTGGTTCATGTTCTCTTCC 1135
 Db 122 GTGCTGGGTTCACTTAATAGAGCTGAACATTAGTCTGCTGGTTCACAGTTCTCTTCT 63
 QY 1136 ATGACCCAGCGCTTCTAATAGAGCTATAACACTCACCGATGCCCCAAGATTCCAT 1191
 Db 62 GAGACCCAGCGCTTCTAATAGAGCTATACGACACTGATGCCCCAAGATTCCAT 7

RESULT 29 AG076758/c

LOCUS AG076758 679 bp DNA linear GSS 03-NOV-2001
 DEFINITION Pan troglodytes DNA, clone: PTB-071A23.R, genomic survey sequence.

AG076758
 ACCESSION AG076758.1 GI:16628560
 VERSION
 KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE AUTHORS

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library PTB

Unpublished

2 (bases 1 to 679)

Totoki, Y., Watanabe, H., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1..679

/organism="Pan troglodytes"

/mol_type="genomic DNA"

FEATURES source

/db_xref="taxon:9598"
 /clone="PTB-071A23.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 32.8%; Score 435.6; DB 9; Length 679;
 Best Local Similarity 86.3%; Pred. No. 2e-120;
 Matches 480; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 133 AATATAATGATAGAGAGAGACCTTCAAAACACTGCACCTCGGGCCTCCTCAGCCAATG 192
 Db 599 AACTCAATGTAGAGCATAGTAGTTTCAGAGCCCGGTCAATTGGGGCTCGTCAGTCAATG 540
 QY 193 GATGCCCTGGACTCTCCCTCTTAGGACCTCTAGCAGCTATAATAATATTTTACTCTCTT 252
 Db 539 GATGCCCTGGGTGCGCCCTCTTAGACCTGTAGCAGCTGTAGGCTGTGCTCTCTCTT 480
 QY 253 TGGACCTGTATCTTCAACTTCTTGTAAAGTTTGTCTTCCAGAAATTGAAGCTGTAAA 312
 Db 479 TGGGCCCTGTATATTCGACCTCTTGTGAGTTTGTCTCTTCAGAAATTAAGGCTGTAAA 420
 QY 313 GCTACAAATAGTTCTTCAATGGAACCCAGATCGATCCATGACTAAATCTACCGTGG 372
 Db 419 GCTACAAATAGTTCTTCAATGGAACCCAGATCGATCCATGACTAAATCTACCGTGG 360
 QY 373 ACCCTGGACCGCTCTAGACTATGCTCTGATGTTAATGACATTGAAGTCAACCCCTCC 432
 Db 359 ACCCTGGGCCAGCTCTAGCCCTTGTCCGATGTTAATGACATCAAGGACCCCTCC 300
 QY 433 CGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCTCAATTCAGTAGGAAGCTAGAG 492
 Db 299 TGAGGAAATCTCAACTGCACAAACCCCTACTACACTCCTCAATTCAGTAGGAAGCTAGAG 240
 QY 493 CAGTTGTGACCAACCTCCCAACAGTACTTGGGTTTTCTGTTGAGAGGTGAGCTGAG 552
 Db 239 CGTCTGCTAGTCAACCTCCAGCAGCACTTGGGTTTTCTGTTGAGAGGTGAGCTGAG 180
 QY 553 AGACAGGACTAGTGGATTTCTTAGCTGCTAAGAAATCCNAAAGCTTANCTGGGAAGGT 612
 Db 179 AGACAGGACTAGTGGATTTCTTAGCCGACTAAGAAATCCNAAAGCTTACCTGGGAAGGT 120
 QY 613 GACCGCATCCATCTTTAAACATGGGGCTTGCAACTTAGCTCACACCGCAATCAGAGA 672
 Db 119 GACCGGTCCACCTTTAAACACGGGCTTGCAACTTAGCTCACACCGCAATCAGAGA 60
 QY 673 GCTCACTAAATGCTA 688
 Db 59 GCTCTCGAGCATGCTA 44

RESULT 30

AG102951

LOCUS

AG102951

DEFINITION

Pan troglodytes DNA, clone: PTB-106G16.F, genomic survey sequence.

AG102951

ACCESSION

AG102951.1

GI:16723468

VERSION

GSS.

KEYWORDS

Pan troglodytes (chimpanzee)

SOURCE

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library PTB

Unpublished

2 (bases 1 to 683)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@gsr.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R Site 1 : SacI

R Site 2 : SacI

Location/Qualifiers

1..683

/organism="Pan troglodytes"

/mol_type="Genomic DNA"

/db_xref="taxon:9598"

/clone="PTB-106G16.F"

/sex="male"

/cell_type="lymphoblast"

/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 32.6%; Score 433.6; DB 9; Length 683;
Best Local Similarity 88.6%; Pred. No. 8.1e-120;
Matches 514; Conservative 0; Mismatches 61; Indels 5; Gaps 4;

QY 547 ACTGAGACGAGCTAGCTGGATTCTTAGCTGACTAAGAAATCCNNAAGCTTANCTGG 606
DB 108 ATTGAGAGAAAGACTAGCTGATTTCTTAGGCCGACTAAGAAATCCCTAAGCTTAGCTGG 167

QY 607 GAAGTGACCGCATCCATCTTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAAT 666
DB 168 GAAGGTGACTGCATCCACCTTTAAACATGGGGCTTGCACTTAGCTCACACCCGACCAAT 227

QY 667 CAGAGGCTCACTAAATCTAATCAGGCAAAACAGGAGTAAAGCAATAGCCAAATCAT 726
DB 228 CAGAGGCTCACTAAATCTAATTAGGCAAAACAGGA-GTAAAGAAACAGCCAAATCAT 286

QY 727 CTATTGCTTGAGACACGCGGGAAGGACAAGGATTGGGATATAAACTCAGGCATTCAAG 786
DB 287 CTATTGCTTGAGACACGCGGGGAGCAAGGATCGGGATATAAACCCAGGCATTGAG 346

QY 787 CAGGACAGAACCCCTTTGGTCCCTCCCATGTATGGAGCTCTGTTTCACTCT 846
DB 347 CCGGCAACGCAACCCCTTTGGTCCCTCC--TTGTAAGGAACCTGTTTCACTCT 404

QY 847 ATTTCACTCTATTAAATCATGCAACTGCACTCTCTGCTCCCTGTTTATGGCTCAAG 906
DB 405 ATTTCACTCTATTAAATCTGCAACTGCTCTCTGCTCCCTGTTTATGGCTCAAG 464

QY 907 CTGAGCTTTTGTTCGCCATCCACCACTGCTG--TTTGCCACCGTCAAGACCCGCTGCTGA 965
DB 465 CTGAGCTTTTCACTTGCTGCTCCACCACTGCTGTTTGTGCTTGTGCTGTTTATGGCTCAAG 524

QY 966 CTTCCATCCCTTTGGATCCAGCAGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025
DB 525 CTTCCATCCCTCGGATCCAGCAGAGTGTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 584

QY 1026 TTGGCACTCCCGATCAGGCTAAGGCTTGCCATGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1085
DB 585 TTGCC-GTTCCGATTGGGCTAAGGCTTGCCATGTTCTGCTGCTGCTGCTGCTGCTGCTGCT 643

QY 1086 TTGTCTTAATAGAACTGAACTGGTCACTGGGTTCCATG 1125
DB 644 TCATCCTAATCGAGCTGAACACTAGTCACTGTGTTCCAGG 683

RESULT 31

AQ381711

LOCUS

DEFINITION AQ381711 611 bp DNA linear GSS 21-MAY-1999

RPC111-165G20.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-165G20,

genomic survey sequence.

AQ381711

AQ381711.1 GI:4352734

GSS.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 611)

AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and

Venter,J.C.

Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

Other GSSs: RPC111-165G20.TV

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@igr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..611

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="GDB:7563139"

/db_xref="taxon:9606"

/clones="RPCI-11-165G20"

/sex="Male"

/cell_type="Lymphocytes"

/clone_lib="RPCI-11"

/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;

RPC111 Human Male BAC Library"

ORIGIN

Query Match 32.4%; Score 430.6; DB 8; Length 611;
Best Local Similarity 86.2%; Pred. No. 6.4e-119;
Matches 517; Conservative 0; Mismatches 62; Indels 21; Gaps 3;

QY 597 GAATCCCNAAAGCTTANCTGGGAAGGTGACCGCATCCATCTTTAAACATGGGGCTTGCAAC 646
DB 22 GAATTCCTAAACCTAGCTGAGAGGTGACCGCATCCACCTTTAAACACGGGGCTTGCAAC 81

QY 647 TTAGCTCACACCCGACCAATC-----AGAGAGCTCACTAAATCTTAATCAGGCAA 697
DB 82 TTAGCTCACACCCGACCAATCAGGTAGTAAAGAGAGCTCACTAAATCTTAATCAGGCAA 141

QY 698 AAACAGGAGGTAAAGCAATAGCCAATCATCTATTGCTTGAGAGCACAGCGGGAAGGACAA 757
DB 142 AAACAGGTGGTAAAGAAACAGCCAATCATCTATTGCTTGAGAGCACAGTGGGAGGACAA 201

QY 758 GGATGGGATATAAACTCAGGCATTCAGCAGCAAGCAACCCCTTTGGTCCCTC 817
DB 202 TGATCGGATATAAAACCCAGGCATTCGAGCCAGCAATGGCTACCTCTTTGGGTCCCTC 261

QY 818 CCATGTATGGAGCTGTGTTTCTCACTCTATTTCATCTATTAAATCATGCAACTG--CA 875
DB 262 TCTTTGTATGGAGCTCTGT-----TTTCACTCTATTAAATCTTGACCTGTACT 311

QY 876 CTCCTTCGTCCGTGTTTTTATGGCTCAAGCTGAGCTTTGTTGCGCATCCACCACTGC 935
DB 312 CTCCTTCGTCCATGTTGTTTATGGCTCAAACTGAGCTTTCACTCGCCATCCCAACTGC 371

QY 936 TGTTTGCCACCGTCAACACCCGCTGTGACTTCCATCCCTTTGGATCCAGCAGATGTC 995


```

Db      372 TGTTCGCCGCTGCACAGATCCGCCACTGATTTCCATCCCTCAGGATCCGGCAGAGTGTC 431
QY      996 CACTGTGCTCTGATCCAGGAGGTACCCATGTCCTCCGATCAGGCTAAAGGCTTGC 1055
Db      432 CGTGTGCTCTGATTCAGGAGGAGCCATTGCCACTCCCAATCAGGCTAAAGGCTNGT 491
QY      1056 CATTGTTCTGCTGATGCTAGTCCCTGGGTTGCTCTTAATAGACTGAACACTGGTCACT 1115
Db      492 AATTGTTCTGCTGATGCTAGTCCCGAGGTTGGTCTTAATTGAGCTGAACGCTAGTCACT 551
QY      1116 GGGTTCATGCTTCTCTTCATCAGCCACCGCTTCTTAATAGACTGAACACTCAACCGCA 1175
Db      552 GGGTTCAGGTAATCTTTCGTCACCCACCGCTTCTTAATAGACTGAACACTCAACCGCA 611

RESULT 32
AG086046/c
LOCUS   Pan troglodytes DNA, clone: PTB-084E11.R, genomic survey sequence.
DEFINITION
ACCESSION AG086046
VERSION   AG086046.1 GI:16637848
KEYWORDS GSS.
SOURCE   Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
AUTHORS  Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
          Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE    BAC end sequences of Library PTB
JOURNAL  Unpublished
REFERENCE
AUTHORS  Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
          Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE    Direct Submission
JOURNAL  Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou, Tsukuba, Ibaraki, 305-8565, Japan
          (E-mail: chimpbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
          Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT  Clones are derived from the chimpanzee BAC library PTB. This BAC end
          was generated during the R&D process and may have higher chance of
          clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector    : pKS145
R.Site 1  : SacI
R.Site 2  : SacI
Location/Qualifiers
1. 736
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clones="PTB-084E11.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES
source
Query Match 32.3%; Score 429.4; DB 9; Length 736;
Best Local Similarity 81.8%; Pred. No. 1.6e-118;
Matches 522; Conservative 0; Mismatches 107; Indels 9; Gaps 2;

QY      670 AGAGCTCACTAAATGCTTAATCAGGCAAAACAGAGTAAAGCAATAGCCAACTCACTA 729
Db      729 AGGGCTCACTAAATGCTTAAT-TGGCGAACAACAGTAAAGCAATAGCCAACTCACTA 671
QY      730 TTGGCTGAGAGCACAGCGGGGAAGGACCAAGGATTTGGGATATATAAATCTAGGCAATCAAGCCA 789
Db      670 TCGCCTGAGAGCACAGTGGGAGGAGCAATAATCGGGATATATAAGCCAGGCAATCAAACTG 611

```

```

QY      790 GCAACAGCAACCCCTTTGGGTCCTCCATTTGATGGAGCTCTGTTTCTACTTAT 849
Db      610 ACAGCGGCTACCTCGTTGGGTCCTCCCTTTGATGGAGCTCTGTTTCTACTTAT 551
QY      850 TCACCTCTATTAATCATGCACTGACACTCTCTCGTCCGCTGTTTATGGCTCAAGCTG 909
Db      550 AGGTCTTGAAGTGCA-----CACTCATCTGGTCCGCTGTTTATGGCTCGAGCTG 499
QY      910 AGCTTTTGTTCGCCATCCACCACTGCTGTTTGCACCGCTCAGACACCCGCTGCTGACTTC 969
Db      498 AGCTTTTGTTCGCCATCCACCACTGCTGTTTGGGGCGCTGTCAGACACCCGCGATGACTTC 439
QY      970 CATCCCTTTGATCCAGCAGAGTGTCCACTGTGCTCTCTGATCCAGCAGAGTACCCATGTC 1029
Db      438 CACCCCTCCGGATCCAGCAGGCTGTGTGCTGCTCTGATCCAGGAGGCGGCCCATGTC 379
QY      1030 CACTCCCGATCAGGCTTAAGGCTTGCCATTGTTCTCTGATGCTAGTCCCTGGGTTTGT 1089
Db      378 CGCTCTCCATCAGGGTGAAGGCTTGGCTTTGTTCTGCAAGATAGTCCCGGTTTCGT 319
QY      1090 CCTAATAGAACTGAACACTGGTTCACCTGGGTTTCATGGTTCTCTTCCATGACCCACGCTT 1149
Db      318 CGTAATGAGCTGAACACTAGTGGCTGGGTTCCACGGATCTCTTCCATGACCCACGCTT 259
QY      1150 CTAATAGACTATAACACTCAGCGATGCGCCAAAGATTCCATTCTTGGTATCTGTGAGG 1209
Db      258 CTAATAGACTATAACACTCAGCGATGCGCTCAAGATTCCATTCTTGGAAATCCGTGAAG 199
QY      1210 CCAGAACCCCGAGTCAAGAGTGGCTGAGGCTTGCACCACTTTGGGAAGTGGCCCACTGCC 1269
Db      198 CCAAGAACTCCAGGTCAGAGAACAGAGGCTTCCCGCCATCTTGGAAATCCGTGCAAG 139
QY      1270 ATTTGCTAGTGGCGCCACCACTCTTGGAGCTGTGG 1307
Db      138 ATCTCGAAGCGGCTTCCCACTCTCTTGGAGCTCTCG 101

RESULT 33
AQ238712 714 bp DNA linear GSS 21-APR-1999
LOCUS   RPC111-64L19.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-64L19,
DEFINITION genomic survey sequence.
ACCESSION AQ238712
VERSION   AQ238712.1 GI:3671003
KEYWORDS GSS.
SOURCE   Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS  Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
          Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.
          Use of human BAC End Sequences for Sequence-Ready Map Building
          Unpublished (1998)
          Other GSSs: RPC111-64L19.TK
          Contact: Mark Adams
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0200
          Fax: 301 838 0208
          Email: mdadams@tigr.org
          Clones are derived from the human BAC library RPCI-11. For BAC
          library availability, please contact Pieter de Jong
          (pieterdejong.med.buffalo.edu). Clones may be purchased from
          BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
          Research Genetics (info@resgen.com). BAC end search page:
          http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
          Seq primer: SP6
          Class: BAC ends.
          Location/Qualifiers
          1. 714
             /organism="Homo sapiens"

FEATURES
source

```

```
/mol_type="genomic DNA"
/db_xref="GDB:7524474"
/db_xref="taxon:9606"
/clone="RPCI-11-64L19"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/notes="Vector: pBCE3.6; Site 1: EcoRI; Site 2: EcoRI;
RPC111 Human Male BAC Library"

ORIGIN
Query Match      32.2%; Score 427.6; DB 8; Length 714;
Best Local Similarity 83.0%; Pred. No. 5.5e-118;
Matches 565; Conservative 0; Mismatches 92; Indels 24; Gaps 6;

QY 589 ATCCNRAAGCCTTANT-GGGAAGGTGACCGATTCATCTTTAAACATGGGCTTGCACACT 647
Db 1 ATTCCTAAGCCTAGCTGGGAAGGTGACACACACCCCTCTTTAAACACAGAGCTTGTAACT 60
QY 648 TAGCTCACACCGCAATC-----AGAGAGCTCACTAAATGCTAATCAGGC-AA 697
Db 61 CAGCTCACACCGGACCTATCAGGTAGTAAGAGAGCTCACTAAATATCAATTAGGCTAA 120
QY 698 AAACAGGAGTAAAGCAAT-AGCCAAATCATCTATTGCTGAGAGCACACGCGGAGGACA 756
Db 121 AAACAGGAGTAAAGAAATATCAATCATCTATCGCTGAGAGCACAGGGGGAGGACA 180
QY 757 AGGATTGGGATATAAATCTCAGGCATTTCAAGCCAGACACAGCAACCCCTTTGGGTCCTCT 816
Db 181 ATGATCGGGATATAAAACCCAGGCAITTTGAGCCAGATCAGGTAAACCTCTTTGGGTCCTCT 240
QY 817 CCCATTGTATGGAGCTCTGTTTTCACTCTATTTCACCTTATTAATCATGCACTG--C 874
Db 241 CACACTGTATGGAGCTCTGT-----TTTCACTCTAATAATCTTGCACTGCACTG 290
QY 875 ACTCTTCTGGTCCGCTGTTTTTATGGCTCAAGCTGAGCTGTTTTGTTGCGCATCCCACTG 934
Db 291 ACTCTTCTGGTCCATGTTGTTCCGCTCAAGCTGAGCTGTTTTGCTCGCGCTCCACCACTG 350
QY 935 CTGTTTGGCACCGTCACAGACCCGCTGTGACTTCGATCCCTTTGGATCCAGCAGAGTGT 994
Db 351 CTGAATGCGGCATTTGCGACCTGCTGCTTGACTTCCACCCCTCCGGATCCGGCAGAGTGT 410
QY 995 CCACTGTGCTCTGATCCAGCAGGTACCCATTGCGACTCCGATCAGCTAAAGGCTTG 1054
Db 411 CCGCTGCACTCTGATCCAGCAGGACCCATTTGCCATCTCCGATCAGGCTAAAGGCTTG 470
QY 1055 CCAATTGTTCTGCTGAGTAAAGTGGCTTGTCTTAATAGAACTGAACACTGGTCACT 1114
Db 471 CCAATTGTTCTGACAGCTAAGTGCTGGTTTCATCTTAATCAGGCTGAACACTGTTCCG 530
QY 1115 TGGGTTCCATGTTCTTCTCCATGACCCAGCGCTTCTAATAGAGCTAACAACCTCACCGC 1174
Db 531 TGGGTTCCACGGTTCTTCTCCATGACTCAACAGCTTCTAATAGAGCTAACAACCTTACCAC 590
QY 1175 ATGGCCCAAGATTCCATTCTTTGGTATCTGTGAGGCAAGACCCAGCTCAGAGAANGT 1234
Db 591 ATGGCCCAAGTTTCAATTGTTGGATCCATGATGCCAAGAACCACCGGTCAGAGAATAA 650
QY 1235 GAGGCTTGCCACCATTTGGGA 1255
Db 651 AAGGCCCGCCCATCTTGGGA 671

RESULT 34
AG108186
LOCUS AG108186 697 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-113D17.R, genomic survey sequence.
ACCESSION AG108186
VERSION AG108186.1 GI:16728704
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 697)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chiebes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. 697
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-113D17.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      31.9%; Score 424.4; DB 9; Length 697;
Best Local Similarity 82.0%; Pred. No. 5.1e-117;
Matches 488; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 669 GAGAGCTCATAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCAATCATCT 728
Db 95 GAGAGCTCATAAATGCTAATTAGGCAAAACAGTAGGTAAAGAAATAGCCAATCATCT 154
QY 729 ATTCCCTGAGAGCACACGGGAGGACAAAGATTTGGATATATAAATCAGGCAATTCAGGCC 788
Db 155 ATTCCCTGAGAGCACACGGGAGGAGCAAGGATCCGGATATAAACCCAGGCATTCAGCT 214
QY 789 AGCAACAGCAACCCCTTTGGGTCCTCCCATTTGATGGAGCTCTGTTTCACTCTAT 848
Db 215 GGCAACGAAACCCCTTTGGGTCCTCCCTTTGATGGAGCTCTGTTTCACTCTAT 274
QY 849 TTCACTCTAATAATCATGCACTGCACTCTCTCGTCCGTTTATTCGGCTCAAGCT 908
Db 275 TTCACTCTAATAATCTTGCAACTGCACTCTCTTCGTTCCGTTGTTGTTAAGCTTGAGCT 334
QY 909 GAGCTTTTGTTCGGCATCCACCACTGCTGTTTGCCACCGTCAAGACCCGCTGCTGACTT 968
Db 335 GAGCTTTTCGTTGCTGTACCAACCGCTGTTTCGCCACCGTCAAGACCCCTCGGTTGGCTT 394
QY 969 CCATCCCTTTGGATCCAGCAGAGTGTCCATGTGCTCTCTGATCCAGCAGAGTACCCATTG 1028
Db 395 ACATCTCTCCGGATCCCGAGGGCGCGCTTGTGCTCTCTGATCCAGCAGCGCTCCCATTTG 454
QY 1029 CCATCCCGATCAGGCTAAAGGCTTGCCATTGTTCTTCGATGGCTAAGTGCTCGGTTTG 1088
Db 455 CCGTTTCCCTATCTCGAATAAAGCCCGCATTTGCTCTTCGACGGCTAACCGTCTGGGTTCA 514
QY 1089 TCCTAATAGAACTGAACACATGGTCACTGGTTCATGTTTCCATGTTCTCTTCCATGACCCAGGCT 1148
Db 515 TCCTAAGCAGAGCTGCCCACTAGTCACTGGGTCCCTCGCTCTCTCTCATGACCCAGGCT 574
QY 1149 TCTAATAGAGTATAAACAACCTCACCGCATGGCCCAAGATTCCATTTCCTTGGTATCTGTGAG 1208
```

```

Db      575 TCTAGTCGAGCTAATACCCACCATATGGTCAAGATTTCCTCTCTGCAACTCGTGA 634
      1209 GCCAAGAACCCAGGTGAGGAAGTGGAGCTTGGCCACCATTTGGGAAGTGGCCC 1263
      635 GCCAAGAACCCAGGTGAGGAAGTGGAGCTTGGCCACCATTTGGGAAGTGGCCC 689

RESULT 35
AG117095
LOCUS      Pan troglodytes DNA, clone: PTB-124N08.F, genomic survey sequence.
DEFINITION
ACCESSION AG117095
VERSION    AG117095.1 GI:16737614
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library PTB
JOURNAL    Unpublished
AUTHORS     2 (bases 1 to 675)
            Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimbases@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT    Clones are derived from the chimpanzee BAC library PTB. This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI.
            Location/Qualifiers
FEATURES     1. .675
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="PTB-124N08.F"
            /sex="male"
            /cell_type="lymphoblast"
            /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match      31.9%; Score 423.8; DB 9; Length 675;
Best Local Similarity 83.9%; Pred. No. 7.7e-117;
Matches 510; Conservative 0; Mismatches 78; Indels 20; Gaps 2;

Qy      671 GAGCTCACTAAATGCTAATCAGGCAAAACAGGAGGTAAAGCAATAGCCCAATCATCTAT 730
      64 GAGCTCACTAAATGCTAATAGGCAACACAGGAGGTAAAGCAATAGCCCAATCATCTGT 123
      731 TGCCTGAGAGACAGCGGGAGGAGGACAGGATTGGGATATAAATCACTGCGCATTCAGGCAG 790
      124 TGCCTGAGAGACAGCGGGAGGAGGACAGGATGATCAGGATATAAATCACTGCGCATTCAGGC 183
      791 CAACAGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTCACTCTATTT 850
      184 CAATGGCAACCCCTTTGGGTCCCTCCCATTTGATGGAGCTCTGTTTCACTCTATTT 243
      851 CACTCTATTAAATCATCACTGACACTCTTCTCGTCCGTGTTTTTTATGGCTCAAGCTGA 910
      244 CACTCTATTAAATCTTGAACACTGACACTCTTCTCGTCCATGTTTGTACGGCTCGAGCTGA 303

Qy      911 GCTTTTGTGGCCATCCACCACTGCTGTGTTTGGCCACCGTCCACAGACCCGCTGCTGATTC 970
      304 GCTTTTGTGGCCATCCACCACTGCTGTGTTTGGCCACCGTCCACAGACCCGCTGCTGATTC 346
      971 ATCCCTTTGGATCCAGCAGAGTGTCCACTGCTCTCTGATCCAGCGAGGTACCATTTGCC 1030
      347 ATCCCTTCAGGATCTGGCAGGGTGTCTCTGTGCTCTGATCCAGCGAGATCCCGTGGC 406
      1031 ACTCCCGATCAGGCTAAAGGCTTGCCATTGTTCTCTGCACTGGCTAAGTGCCTGGGTTTGT 1090
      407 ACTCCCATCTGCTGAAGGCTTGCCATTGTTCTCTGCACTAAGTGCCTGGGTTTCACT 466
      1091 CTAATAGAACTGAACACTGCTCACTGGGTTCATGTTCTCTTCATGACCCACGGCTTC 1150
      467 CTAATCGAGCTGAACACTAGTCACTGGGTTCACCGGTTCTCTTCCTGACCCACGGCTTC 526
      1151 TAATAGAGCTATAACACTCACCGCATGGCCAGATTCATTCCTTGGTATCTGTGAGGC 1210
      527 TAATAGAGCTATAACACTCACCGCATGGCCAGATTCATTCCTTGGTATCTGTGAGGC 586
      1211 CAAGAACCCCGCTCAGGAANGTGA---GGCTTGCACCATTTGGGAAGTGGCCCACTG 1267
      587 CAAGAAAGCTTGCACCATCTTGGAGTGGCTGCGCACCATCTTGGAGTGGTTCGCAC 646
      1268 CCATTTTG 1275
      647 CATCTTTG 654

RESULT 36
CB990962
LOCUS      797 bp mRNA linear EST 01-MAY-2003
DEFINITION AGENCOURT 13518713 NIH MGC.148 Homo sapiens cDNA clone
            IMAGE:30334786 5', mRNA sequence.
ACCESSION CB990962
VERSION    CB990962.1 GI:30285482
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 797)
            NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM355 row: j column: 11
            High quality sequence stop: 661.
            Location/Qualifiers
FEATURES     1. .797
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:30334786"
            /tissue_type="pre-eclampsia placenta"
            /lab_host="DH10B Tona"
            /clone_lib="NIH MGC 148"
            /notes="Organ: placenta;
            all-xhoI: Site 2: BamH; Library is oligo-dT primed and
            directionally cloned using primer
            5'-TTTCTTTTCTTTTCTTTT-3', size-selected for average insert
            size 2.3 kb and normalized to R01 5. This is a primary
            library enriched for full-length clones and constructed
            using the Cap-trapper method (Carninci, in preparation)."

```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 826)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM359 row: k column: 16
High quality sequence stop: 585.

FEATURES

1..826
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3036351"
/tissue_type="pre-eclampitic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: Placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 31.8%; Score 422.2; DB 6; Length 826;
Best Local Similarity 88.5%; Pred. No. 2.5e-116;
Matches 469; Conservative 0; Mismatches 59; Indels 2; Gaps 1;
QY 800 CCCCCTTTGGTCCCTCCCATGTATGGAGCTCTGTTTCACTCTATTCTACTCTATT 859
DB 34 CCCCCTTTGGTCCCTCCCTTTGTATGGAG--CTGTTTTCATGCTATTCTACTCTATT 91
QY 860 AAATCATGCAACTGCACTCTCTGTCGCTGTTTATGGCTCAAGCTGAGCTTTTGT 919
DB 92 AAATCTTGAACATGCACTCTCTGTCGCTGTTTATGGCTCAAGCTGAGCTTTTGT 151
QY 920 CGGCATCCACCACCTGCTGTTTGGCCACCGTCACAGACCCGCTGCTGATCCATCCCTTTG 979
DB 152 CACCGTCCACCACTGCTGTTTGGCCACCGACCGCTGCTGATCCATCCCTTTG 211
QY 980 GATCCAGAGAGTGTCCATGTTTCTGTCGCTGCTGTTTATGGCTCAAGCTGAGCTTTTGT 1039
DB 212 GATCCTGACAGGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 271
QY 1040 CAGGCTAAAGCTTGGCAATGTTTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099
DB 272 TGGGCTAAAGCTTGGCAATGTTTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 331
QY 1100 CTGAACACTGCTCACTGGGTTCCATGTTTCTCTTCCATGACCCACGCTTCTTAATAGAGC 1159
DB 332 CTGAACACTAGTCACTGGGTTCCATGTTTCTCTTCTGTCGCTGCTGCTGCTGCTGCTGCTGCT 391
QY 1160 TATAAACAATCAACGATGCGCCAAAGATTCATTCCTGTTGGTATCTGTGAGGCGCAAGACCC 1219
DB 392 TATAACAATCTACCAATGCGCCAAAGATTCATTCCTGTTGGTATCTGTGAGGCGCAAGACCC 451
QY 1220 CAGGTCAAGAGTGTGAGGCTTGGCCACCATTTGGGAAGTGGCCACCTGCTGCTGCTGCTGCTGCT 1279

Db 452 CAGGTCAAGAGTGTGAGGCTTGGCCACCATTTTGAAGCGGCTGCTTACCATCTTTGGAAG 511
QY 1280 CGGCCCAACCACTCTTTGGAGCTGTGGAGCAAGGATCCCCAGTAACA 1329
Db 512 TGGTTCAACCACTCTTTGGAGCTGTGTGAGCAAGGACCCCGGTAAACA 561

RESULT 39

CB996320
LOCUS
DEFINITION
IMAGE:3033232 5', mRNA sequence.
CB996320
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB996320 837 bp mRNA linear EST 01-MAY-2003
AGENCOURT 13617032 NIH_MGC_148 Homo sapiens cDNA clone
IMAGE:3033232 5', mRNA sequence.
CB996320
CB996320.1 GI:30290840
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 837)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM348 row: p column: 01
High quality sequence stop: 611.

REFERENCE

1 (bases 1 to 837)
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM348 row: p column: 01
High quality sequence stop: 611.

FEATURES

Location/Qualifiers
1..837
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3033232"
/tissue_type="pre-eclampitic placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_148"
/note="Organ: Placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Library is oligo-dT primed and
directionally cloned using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
size 2.3 kb and normalized to ROT 5. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHGRI,
National Institutes of Health). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 31.8%; Score 422.2; DB 6; Length 837;
Best Local Similarity 88.5%; Pred. No. 2.5e-116;
Matches 469; Conservative 0; Mismatches 59; Indels 2; Gaps 1;
QY 800 CCCCCTTTGGTCCCTCCCATGTATGGAGCTCTGTTTCACTCTATTCTACTCTATT 859
DB 34 CCCCCTTTGGTCCCTTCCCTTTGTATGGAG--CTGTTTTCATGCTATTCTACTCTATT 91
QY 860 AAATCATGCAACTGCACTCTCTGTCGCTGTTTATGGCTCAAGCTGAGCTTTTGT 919
DB 92 AAATCTTGAACATGCACTCTCTGTCGCTGTTTATGGCTCAAGCTGAGCTTTTGT 151
QY 920 CGGCATCCACCACCTGCTGTTTGGCCACCGTCACAGACCCGCTGCTGATCCATCCCTTTG 979
DB 152 CACCGTCCACCACTGCTGTTTGGCCACCGACCGCTGCTGATCCATCCCTTTG 211
QY 980 GATCCAGAGAGTGTCCATGTTTCTGTCGCTGCTGTTTATGGCTCAAGCTGAGCTTTTGT 1039

[illegible]

RESULT 40	
CB994833	
LOCUS	852 bp mRNA linear EST 01-MAY-2003
DEFINITION	ACENGCOURT13633532 NIH MGC 148 Homo sapiens cDNA clone IMAGE:30339155 5', mRNA sequence.
ACCESSION	CB994833
VERSION	CB994833.1 GI:302899353
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 852)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgapbs-1@mail.nih.gov
	Tissue Procurement: Dr. Stefan Hansson
	cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA sequencing by: Agencourt Bioscience Corporation
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

FEATURES
source

ORIGIN